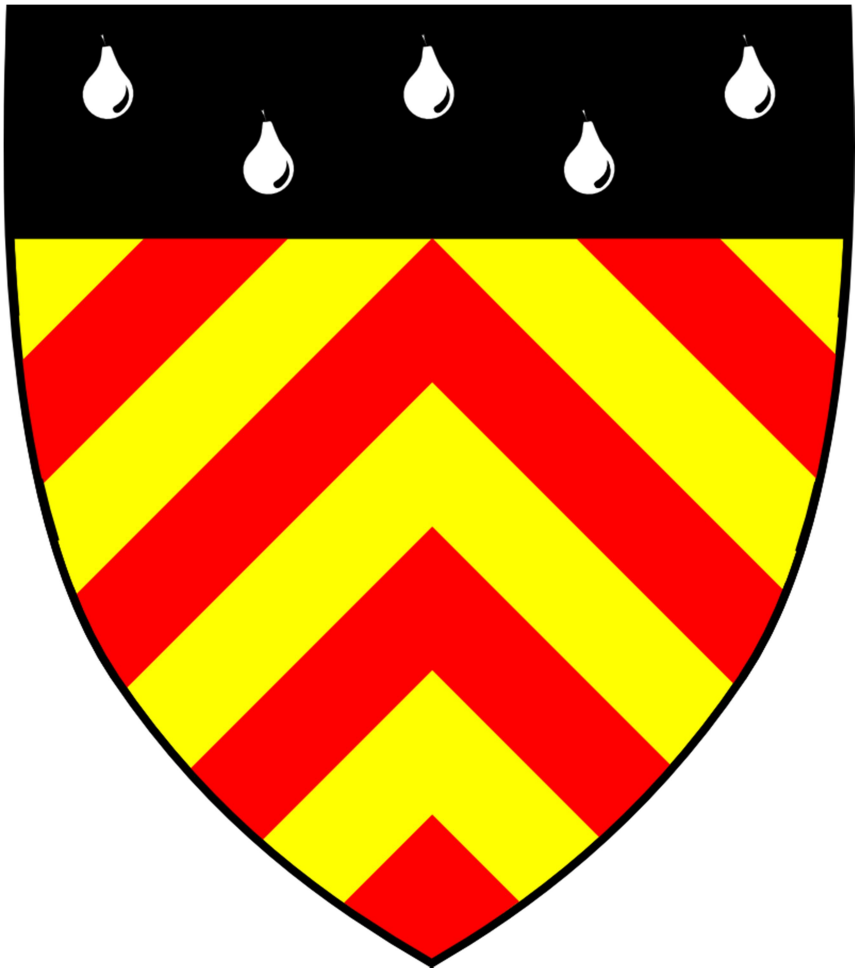
The ancestry and function of cytochrome c_{6A}

Barnaby Andrew Slater



Clare Hall College, University of Cambridge

Submitted September 2019

This dissertation is submitted for the degree of Doctor of Philosophy

Preface

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

This thesis is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

This thesis does not exceed the prescribed word limit for the relevant Degree Committee.

Barnaby Slater

Acknowledgements

Firstly, I would like to thank Professor Chris Howe for all his supervision and guidance throughout my studies and thesis writing. I would also like to thank the rest of the Howe lab for their general support, guidance and thoughts on results during my research project. I give particular thanks to Isabel Nimmo for her words of encouragement and wisdom throughout, as well as teaching me the ways of the FPLC and *E. coli* expression of haemoproteins. Also, from the Howe lab I would like to give special thanks to Stephen Rowden and Ravendran Vasudevan for their assistance in processing and interpreting the RNA sequencing results. I would like to thank Adrian Barbrook for the supplementation of many constructs for expression in *E. coli* that were vital for my research, and Saul Purton and Aleix Gorchs Rovira for the provision of *C.reinhardtii* strains. Finally, from his time at the Howe lab I would like to thank Hamish Symington for being an exceptionally helpful and efficient student, and his contribution to my research.

I also express gratitude to my Graduate Thesis Panel Ross Waller, James Locke and Paul Dupree for their incredibly logical and useful advice during the first two years of my PhD. Aron Ferenczi and Ravendran Vasudevan of the Molnar lab in Edinburgh created and screened the CrispR-cpf1 mutant *C. reinhardtii* that was utilised in a majority of my research and thesis, and for that I am very grateful. I would also like to thank Dr Katherine Stott for her assistance with circular dichroism, and access to her CD spectrometer. Similarly I would like to thank Marko Hyvonen for the use of his size exclusion FPLC column. From the Plant Sciences department I would like to thank Gareth Steed and

Howard Griffiths, for the use of the chlorophyll fluorescence imager and the latter's expertise on carbon concentrating mechanisms.

I would like to give a special thanks to my nuclear family: Andrew, Joanna and Rebecca Slater for their continuing emotional and financial support throughout my studies. I would also like to extend my gratitude towards Lewis Bligh, Daniel Heap and Joseph Horne, for helping me through the ups and downs of my degree. Finally, I would like to thank Jessica Law, for keeping me sane.

This study, alongside my living costs during the study, was funded by the BBSRC.

Abstract

Cytochrome c_{6A} is a homologue of cytochrome c_6 found in eukaryotic green algae and higher plants. However it is thought to perform a different function from cytochrome c_6. Two current hypotheses exist for this function: that cytochrome c_{6A} acts as a 'safety valve', providing an alternative path for electron flow through the photosynthetic electron transport chain and thus alleviating reactive oxygen species production; that cytochrome c_{6A} has a signalling role where it could sense high light stress and convey this signal to the nucleus to affect gene expression.

This study aims to provide insight into the ancestry and function of cytochrome c_{6A}. To gain a clearer understanding of the evolutionary history of cytochrome c_{6A} and cyanobacterial homologues cytochrome c_{6B} and c_{6C}, phylogenetic analysis was performed on peptide sequences of these proteins as found in a wide range of photosynthetic organisms. These data were used to refine the model of the evolutionary history of the cytochrome c_6 family. This study also found evidence that cytochrome c_{6B} and c_{6C} are in fact orthologues with a similar function.

Chlamydomonas reinhardtii is a model organism for eukaryotic photosynthesis, as its unicellularity provides many advantages over land plants. To determine if cytochrome c_{6A} is involved in high light stress in *C. reinhardtii*, cytochrome c_{6A} knockout and knockdown mutant lines of *C. reinhardtii* had been previously produced using CrispR-cpf,1 and are characterised in this study. The mutant lines demonstrated potential growth retardation under high or fluctuating light stress, as well as singlet oxygen stress, which was more noticeable under mixotrophic conditions. Chlorophyll fluorescence analysis of the mutant lines established a link between cytochrome c_{6A} and NPQ. As cytochrome c_{6A} is believed to be located in the thylakoid lumen, and initial NPQ is triggered by a change in luminal pH, circular dichroism was used to determine changes in secondary structure of purified cytochrome c_{6A} over a pH range of 2-7. No significant change in structure was observed, but cytochrome c_{6A} did maintain structural integrity even at pH 2. Finally, the transcriptome of the cytochrome c_{6A} knock out line was compared to the background strain under standard and high light conditions through RNAseq. Many photoprotective, motility and CCM genes were differentially regulated under high light stress, but when cytochrome c_{6A} was knocked out these regulations were diminished. Therefore cytochrome c_{6A} has been proposed as a signalling molecule that functions in CCM, NPQ and motility under photosynthetic stress conditions.

Table of Contents

List of abbreviations

2PG: 2-phosphoglycerate

3PG: 3-phosphoglycerate

ATP: Adenosine triphosphate

BSA: Bovine serum albumin

CCM: Carbon concentration mechanism

CD: Circular dichroism

CDK: Cyclin-dependent kinase

CDS: Coding sequence

CEF: Cyclic electron flow

CVDE: Chlorophycean violoxanthin de-epoxidase

DBMIB: Dibromothymoquinone

DCMU: 3-(3,4-dichlorophenyl)-1,1-dimethylurea

DNA: Deoxyribonucleic acid

EGF: Epidermal growth factor

EST: Expression sequence tag

FAP: Flagellar associated protein

FPLC: Fast protein liquid chromatography

GDP: Guanosine diphosphate

GPXH: Glutathione peroxidase homologue

GST: Glutathione-S-Transferase

HCl: Hydrochloric acid

HL: High light

LB: Lysogeny broth

LHC: Light harvesting complex

LIP: Loop insertion peptide

LL: Low light

MAPK(KK): Mitogen-activated protein kinase

MeOH: Methanol

ML: Maximum-Likelihood

NADP: Nicotinamide adenine dinucleotide phosphate

NJ: Neighbor-Joining

NPQ: Non-photochemical quenching

NPQf: Fast-relaxing NPQ

NPQs: Slow-relaxing NPQ

OMD: Overall mean distance

ORF: Open Reading Frame

PCR: Polymerase chain reaction

PETC: Photosynthetic electron transport chain

PSI: Photosystem I

PSII: Photosystem II

RB: Rose Bengal

RET: Resonance energy transfer

RES: Reactive electrophilic species

RIN: RNA integrity number

RNA: Ribonucleic acid

RNAseq: RNA sequencing

ROS: Reactive oxygenic species

SDS-PAGE: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SL: Standard light

SP: Saturating pulse

TAP: Tris Acetate Phosphate

TP: Tris Phosphate

UTR: Untranslated region

WGS: Whole genome sequencing

1. Introduction

1.1. *Chlamydomonas reinhardtii* is a model organism for eukaryotic photosynthesis

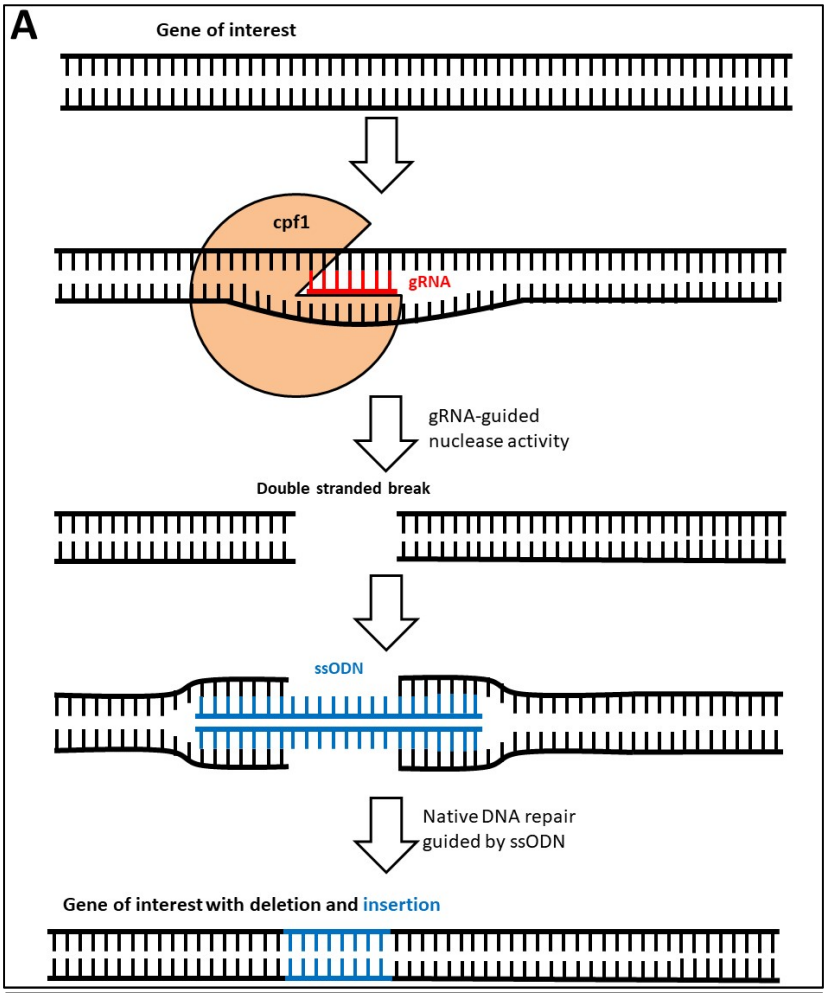
Chlamydomonas reinhardtii is an unicellular eukaryotic freshwater green alga. It consists of a 10 μm diameter spheroid-shaped cell encapsulated in a crystalline glycoprotein cell wall, with two flagella protruding from the cell's apex, and one large cup-shaped chloroplast (Sager and Palade, 1957; Randall, 1969; Domozych *et al*., 2012). *C. reinhardtii* also has a pyrenoid, the centre for CO_2 fixation and the site of carbon-concentration mechanisms, and an eyespot used for light detection and phototaxis (Goodenough, 1970; Boscov and Feinleib, 1979; Melkonian and Robenek, 1980).

C. reinhardtii is frequently used as a model organism for eukaryotic photosynthesis. This is because it performs photosynthesis in a chloroplast by a very similar mechanism to that in higher plants, but has the advantages of a unicellular organism. *C. reinhardtii* can be grown on solid or liquid media, has a simple life cycle, and has a relatively fast growth rate compared to higher plants, meaning it is easier to grow and maintain in the laboratory. The alga has a fully sequenced haploid genome (Merchant *et al.*, 2007), making it easier to manipulate, and methods of genetic modification have been published for all three genomes: nuclear, chloroplast and mitochondria (Rochaix, 1995; Stauber and Hippler, 2004). *C. reinhardtii* is also capable of growing either photoautotrophically or mixotrophically with an exogenous carbon source, such as acetate (Sager and Granick, 1953). This means that the alga can be grown in the dark, and allows for more effective analysis of mutant lines deficient in photosynthesis. As well as being a very useful model organism for eukaryotic photosynthesis, *C. reinhardti* has also been used in biofuel production and eukaryotic expression systems for pharmaceutical production (Scranton *et al*., 2015; Yan *et al*., 2016). Therefore, understanding how *C. reinhardtii* performs photosynthesis and protects the photosynthetic machinery from light stress will benefit both fundamental plant science and bioproduction.

1.1.1. A CRISPR/Cpf1 system can be used to genetically modify *Chlamydomonas reinhardtii*

Recently a method using the CRISPR/Cpf1 targeted nuclease mechanism has been developed for deleting regions of the *C. reinhardtii* genome (Ferenczi *et al.*, 2017; Figure 1.1), allowing for more efficient gene modification than through homologous recombination. The Cpf1 enzyme is guided to the DNA target through guide RNAs (gRNAs), where a double stranded cut is made (figure 1.1A). The native DNA repair mechanism of *C. reinhardtii* is then recruited to the double stranded break and repairs it. The combination of this CRISPR/Cpf1 system with single stranded oligodeoxynucleotides

(ssODN) as DNA repair templates has proved much more efficient for *C. reinhardtii* gene modification than the similar CRISPR/Cas9 system (Jiang *et al*., 2014). Inclusion of ssODNs as DNA repair templates can guide the repair mechanism to insert a foreign sequence of DNA in the double stranded break, which could be an inserted marker gene or a stop codon repeat to interrupt correct translation. If two gRNAs targeting separate regions of a gene are used simultaneously, Cpf1 will make double stranded breaks in both regions. Then an ssODN that has complementary linkers to the sequences flanking the gRNA target regions will make the DNA repair mechanism connect the flanking sequences at the exclusion of the region between the gRNA targets (figure 1.1B). This combined with an inserted stop codon repeat sequence allows for an efficient knockout mutation of a specific gene in *C. reinhardtii*.


A
Gene of interest
cpf1
gRNA
gRNA-guided
nuclease activity
Double stranded break
ssODN
Native DNA repair
guided by ssODN
Gene of interest with deletion and insertion

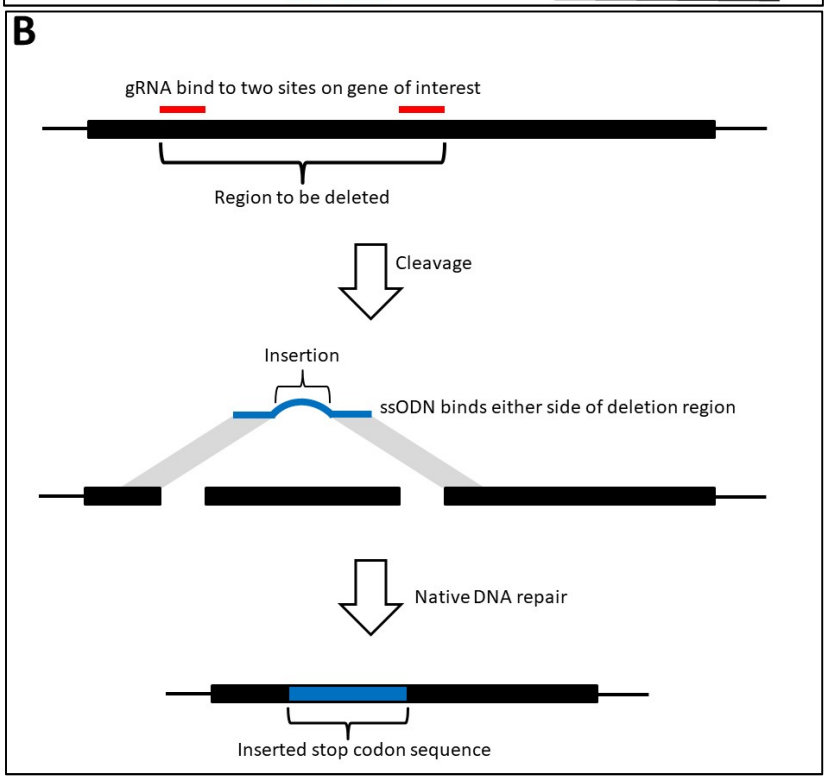

B
gRNA bind to two sites on gene of interest
Region to be deleted
Cleavage
Insertion
ssODN binds either side of deletion region
Native DNA repair
Inserted stop codon sequence

Figure 1.1: Method of gene manipulation using CRISPR/Cpf1 in C. reinhardtii. A) Mechanism for a single guide RNA (gRNA, red)-guided deletion of a small region of a gene followed by insertion guided by a single stranded oligodeoxynucleotide (ssODN, blue). B) Mechanism for larger-scale deletion of a region of a target gene (black box), in which two double stranded breaks are induced with separate gRNAs. The ssODN that guides the native repair mechanism then is complementary to the regions flanking those that binded the gRNA (grey parallelograms represent complementarity). This removes an entire region of the gene of interest and replaces it with an inserted stop codon sequence.

1.2. Photosynthesis and photoprotection in *Chlamydomonas reinhardtii*

1.2.1. Photosynthesis in *Chlamydomonas reinhardtii*

C. reinhardtii uses light energy to produce reducing equivalents and drive ATP synthesis through the photosynthetic electron transport chain (PETC; figure 1.2). Light energy in the form of photons cause electrons in chlorophyll and other pigments in the light harvesting complex to enter an excited state. When the electrons return to the ground state, the energy is transferred to adjacent pigments by resonance energy transfer (RET). Eventually, the chlorophylls in the PSII reaction centre receive this energy, and the excited electron can be transferred to Q_A via phaeophytin. The oxidised reaction centre chlorophyll can then be reduced by electrons transferred from the oxygen evolving complex after water oxidation. Reduced Q_A will transfer the electron to plastoquinone, which then diffuses into the thylakoid membrane (plastoquinone pool). The plastoquinol in turn reduces cytochrome *f* in the cytochrome b_6f complex by undergoing the Q cycle. Cytochrome *f* then reduces plastocyanin, which is located in the thylakoid lumen, and transfers the electron to the reaction centre of PSI. In a similar way to PSII, RET from light harvesting complexes allows charge separation at the PSI reaction centre, and the electron is passed through FeS clusters to ferredoxin. Ferredoxin can then be used to reduce $NADP^+$, catalysed by ferredoxin-$NADP^+$ reductase. During the electron transfer process, photohydrolysis and the Q cycle release protons into the thylakoid lumen, and plastoquinone and $NADP^+$ reduction removes protons from the stroma. Cytochrome b_6f also pumps protons across the thylakoid membrane into the lumen. This collectively produces an electrochemical gradient across the thylakoid membrane, which can drive ATP synthesis. Thus the PETC uses light energy to drive the production of ATP and reducing equivalents, which can be used in other processes, particularly the Calvin cycle and CO_2 fixation. It is important to note that in *C. reinhardtii* and many other prokaryotic and eukaryotic algae, cytochrome c_6 can replace plastocyanin functionally when copper levels in the environment are limiting (Wood, 1978).

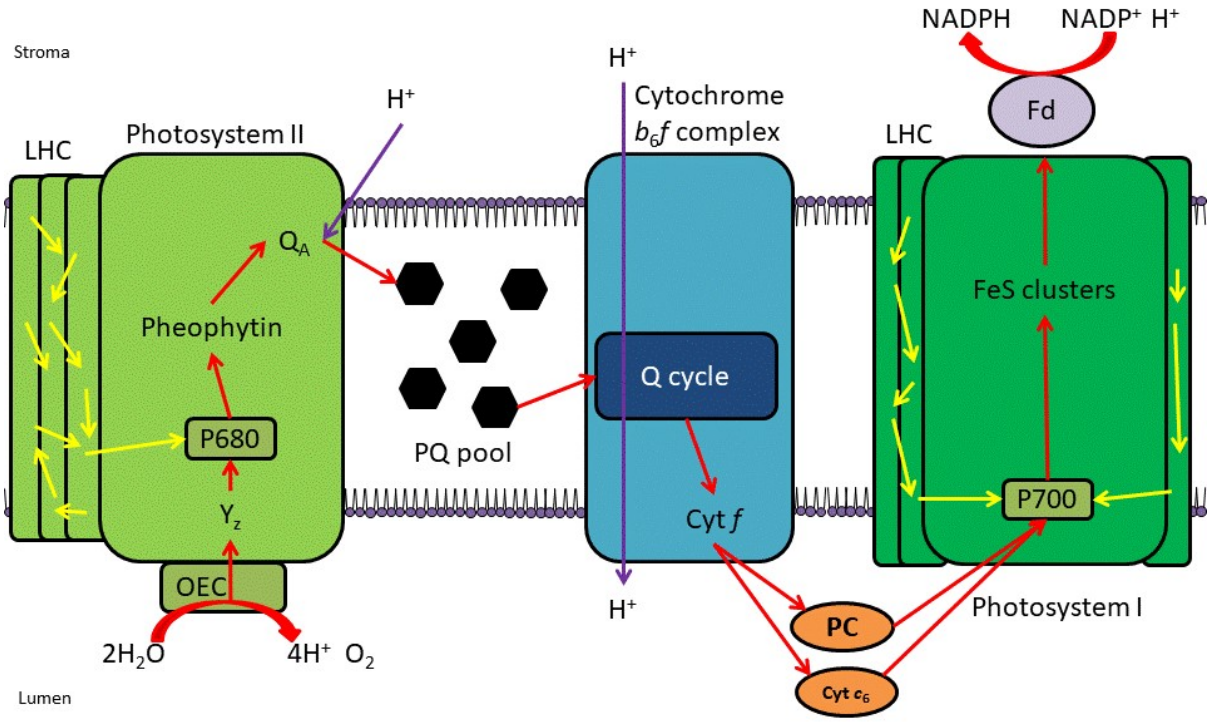


Figure 1.2: The linear photosynthetic electron transport chain. The red arrows indicate electron flow, yellow arrows indicate resonance energy transfer and purple arrows proton movement. Abbreviations are as follows: LHC, light harvesting complex; OEC, oxygen evolving complex; PQ, plastoquinone; cyt, cytochrome; FeS, iron-sulphur; PC, plastocyanin; Fd, ferredoxin. P680 and P700 represent the PSII and PSI reaction centres respectively.

1.2.2. *C. reinhardtii* has multiple systems in place to deal with light stress

1.2.2.1. High light exposure causes stress in C. reinhardtii

There is a limit to how much absorbed light energy can be used to drive photosynthesis. If the rate of photons being absorbed exceeds this limit, then this can lead to the production of harmful chemicals called reactive oxygenic species (ROS). Singlet oxygen is one such ROS predominantly formed at PSII (Papageorgiou, 2014). The bottleneck in the PETC usually occurs at the Q cycle in cytochrome b_6f (Papageorgiou, 2007), which creates an over-reduced plastoquinone pool. The backup means that P680 spends most of the time in the reduced state, and the excited chlorophyll in the light harvesting complex cannot transfer the energy to the reaction centre. The longer the chlorophyll is kept in the excited state, the more likely it is to enter the triplet state. Triplet state chlorophyll can then react with molecular oxygen to produce singlet oxygen, which is able to react with and thus damage various biomolecules promiscuously (Krieger-Liszkay, 2005; Ledford *et al.*, 2007). Therefore

a key aspect of photoprotection is to alleviate singlet oxygen, through returning the ROS to molecular oxygen, through quenching high energy chlorophyll before it can reach the triplet state, or through decreasing the number of high energy chlorophyll being formed.

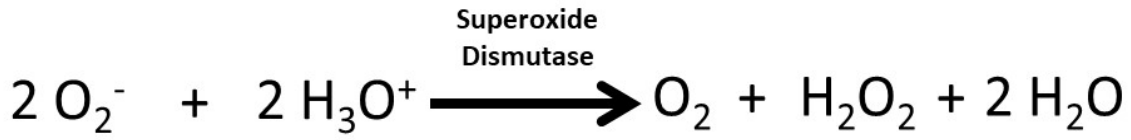


Figure 1.3: Equation for the reaction catalysed by superoxide dismutase.

Superoxide is predominantly produced at PSI as a result of molecular oxygen photoreduction (Asada *et al*., 1974). The superoxide is then disproportionated into hydrogen peroxide, catalysed by superoxide dismutase (Bowler *et al*., 1994; figure 1.3). Similarly to singlet oxygen, superoxide and hydrogen peroxide are very damaging to the photosynthetic apparatus, and therefore enzymes such as ascorbate peroxidase reduce hydrogen peroxide to water as part of the ascorbate-glutathione cycle (Asada, 1992; Figure 1.4). This cycle, through the redox reactions of several intermediates including glutathione, effectively allows enzyme-controlled reduction of hydrogen peroxide to water using reducing equivalents NADH or NADPH. Protection of the photosynthetic machinery against attack from ROS is therefore essential to the functioning of *C. reinhardtii*.

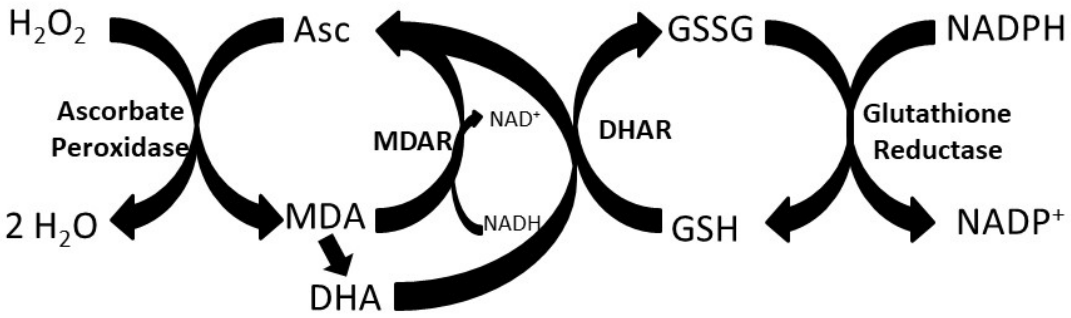


Figure 1.4: The ascorbate-glutathione cycle of hydrogen peroxide oxidation. Ascorbate (Asc) is reduced to monodehydroascorbate (MDA). MDA can in turn be reduced to ascorbate through NADH oxidation, catalysed by monodehydroascorbate reductase (MDAR). MDA can also disproportionate to dehydroascorbate (DHA), which can be reduced to ascorbate through glutathione (GSH) oxidation, catalysed by dehydroascorbate reductase (DHAR). Oxidised glutathione (GSSG) can then be reduced back to glutathione by the activity of glutathione reductase, which in turn oxidises NADPH.

1.2.2.2. Short term strategies of photoprotection

Photoprotection in *C. reinhardtii* can be categorised into short term responses (from seconds to minutes) and long term responses (hours to days; Erickson *et al*., 2015). Typically, short term responses allow for survival in fluctuating light conditions, such as changes in weather, diurnal cycle or depth in freshwater habitats. The main objective of photoprotection is to prevent chlorophyll from being in the excited state for too long, as this increases the chance of entering the triplet state and ROS production. This can be achieved by a combination of two methods: photochemical quenching, where chlorophyll is quenched by energy transfer through RET through to the reaction centres and photosynthesis; and non-photochemical quenching (NPQ), where excited chlorophyll are brought back to the ground state and the energy is dissipated as heat (Müller *et al.*, 2001).

Photochemical quenching functions under the principle that increased electron flow from PSII will increase the rate of RET from chlorophyll and thus a decrease in ROS production. In the short term this can be achieved through the use of alternative electron pathways which create a 'safety valve' effect (figure 1.5). One such pathway, denoted cyclic electron flow (CEF), involves the reduction of the plastoquinone pool using electrons from charge separation at PSI (Finazzi *et al.*, 1999). This is achieved through the transfer of electrons from NADPH to plastoquinone via an NADPH-plastoquinone reductase, which was found to be NDA2 in *C. reinhardtii* (Jans *et al.*, 2008). Interestingly, a ferredoxin-plastoquinone oxidoreductase complex called PGR5-PGRL1 has also been found in *C. reinhardtii*, which would allow CEF to occur by ferredoxin directly reducing plastoquinone (Hertle *et al.*, 2013). Therefore this suggests two possible CEF pathways in *C. reinhardtii*. These cycles allow proton pumping across the membrane without water splitting or reducing equivalent production, and therefore are used to modulate the pH gradient across the thylakoid membrane, which in turn will affect NPQ. In fact, mutants that cannot undergo CEF were found to have increased sensitivity to high light intensity and have decreased NPQ (Tolleter *et al*., 2011; Johnson *et al.*, 2014). However, it should be noted that CEF is more than a photoprotective mechanism, as there may be other contexts where ATP is required without the need for NADPH. Another alternative electron pathway that can contribute to photoprotection is the production of water through oxygen reduction, known as the water-water cycle (Asada, 2000; figure 1.5). Flavodiiron proteins have been found in *C. reinhardtii*, and have been shown to play a role in molecular oxygen reduction using NADPH (Dang *et al.*, 2014; Chaux *et al.*, 2017; Kosourov, *et al.*, 2018). The superoxide conversion to hydrogen peroxide and water mentioned in 1.2.2.1 would also count as a water-water cycle pathway. These pathways, despite seeming futile cycles, allow for the continued transfer of

electrons through the PETC under high light conditions, which decreases the subsequent production of singlet oxygen at PSII.

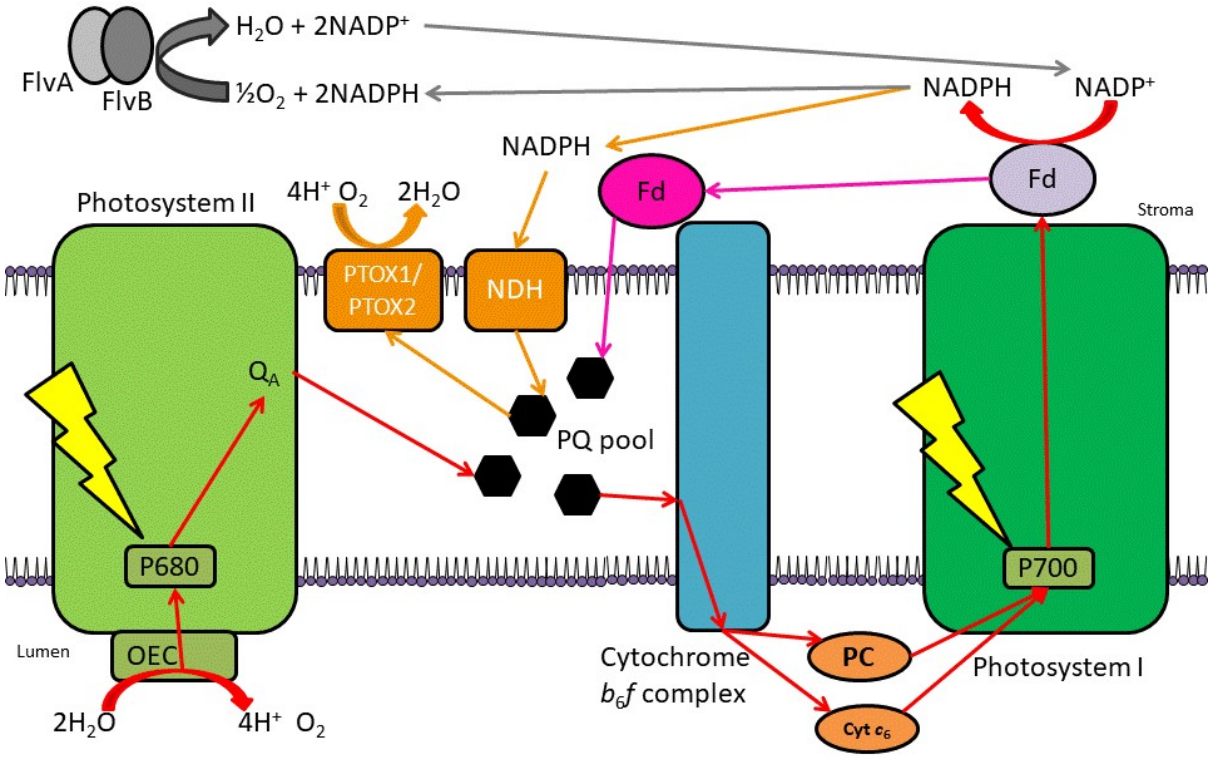


Figure 1.5: The photosynthetic electron transport chain with alternative electron pathways included. Red arrows represent linear electron flow, pink arrows represent electron transfer exclusive to cyclic electron flow, orange arrows represent electron transfer involved in chlororespiration, and grey arrows represent the water-water cycle. Lightning symbols represent energy from photons reaching the P680 and P700 reaction centres. Abbreviations are: OEC oxygen evolving complex; Q quinone; PQ plastoquinone; PC plastocyanin; cyt cytochrome; Fd ferredoxin; Flv flavodiiron; PTOX plastid terminal oxidase; NDH NADPH dehydrogenase.

Respiration can also be a useful way of providing photochemical quenching. As mitochondrial respiration uses reducing equivalents to drive oxidative phosphorylation, increased respiration can help alleviate high light stress through acting as a sink for electrons and driving photosynthesis. This was shown by an observed link between increased mitochondrial activity and high light stress in *C. reinhardtii* (Davis *et al*., 2013), and has been also shown to occur in cyanobacteria (Lea-Smith *et al.*, 2013). Davis *et al.* (2013) also found an increase in amino acid production under high light stress, suggesting that amino acid synthesis is also used as a sink for reducing equivalents in order to increase photosynthesis and photochemical quenching. A form of respiration has also been observed to occur in the chloroplast, known as chlororespiration (Bennoun, 1982). This is where NADPH is

used to reduce molecular oxygen via plastoquinone and plastid terminal oxidases (PTOX1 and PTOX2 in *C. reinhardtii*; Cournac *et al*., 2000; Houille-Vernes *et al*., 2011). The electrons are passed to plastoquinone through NDA2 activity, and then plastoquinol reduces oxygen at the plastid terminal oxidases. This system can act as a safety valve for photosynthesis, as the terminal oxidases can alleviate an over-reduced plastoquinone pool, and the NADPH-plastoquinone oxidoreductase and alleviate an under-reduced plastoquinone pool (figure 1.5). This pathway is particularly useful when PSI activity or expression is low.

Non-photochemical quenching can be separated itself into several processes to allow chlorophyll quenching at PSII. The fastest NPQ response is called rapidly reversible NPQ (qE). This response is activated by the acidification of the lumen (Briantais *et al.*, 1979), which will occur when the electron transfer rate through the PETC exceeds that of proton gradient dissipation through ATP synthesis. The activation of qE by the pH difference between stroma and lumen was shown with a study where *C. reinhardtii* exposed to the ionophore nigericin lacked an NPQ response (Niyogi *et al*., 1997a). A low pH in the thylakoid lumen can in itself be detrimental to photosynthesis, as it causes damage to the oxygen evolving complex (Virgin *et al.*, 1988). The low pH has been shown to be detected by the C-terminus of the LHCSR3 protein, which protrudes into the thylakoid lumen (Peers *et al*., 2009; Ligouri *et al*., 2013). At low pH, LHCSR3 bound to chlorophyll and xanthophylls can quench chlorophyll at the light harvesting complexes (Bonente *et al*., 2011). A different system is used by higher plants, where PsbS is the detector of lumen acidification, and activates qE through activating pigment-bound LhcB proteins (Li *et al*., 2004; Niyogi *et al*., 2001). A lutein radical cation has also been found to be involved in qE in both green eukaryotic algae and higher plants, and is also thought to quench excited chlorophyll (Anwaruzzaman *et al.*, 2004; Bonente *et al*., 2011). Inactivation of both lutein and zeaxanthin biosynthesis causes qE to be completely eliminated in *C. reinhardtii* (Niyogi *et al.*, 1997b; Amanarth *et al*., 2012). This combined with the fact that either knock out on its own only impairs qE but does not lead to its elimination suggests that two mechanisms exist for qE, one involving lutein and one involving LHCSR3 and zeaxanthin.

After the initial qE NPQ response comes a second phase of zeaxanthin-dependent NPQ (qZ). During this, zeaxanthin levels accumulate, allowing increased quenching of excited chlorophyll. This was shown using a mutant line of *C. reinhardtii* deficient in the chlorophyte-specific violoxanthin de-epoxidase (CVDE) function. qE still occurred in this line, but a second phase (now identified as qZ) was diminished (Niyogi *et al*., 1997a; Li *et al.*, 2016). CVDE catalyses the de-epoxidisation of violaxanthin into zeaxanthin in the xanthophyll cycle in response to high light (Hieber *et al*., 2000; Figure 1.6). In low light the zeaxanthin is epoxidised back to violaxanthin by zeaxanthin epoxidase

(Baroli *et al*., 2003). A mutant line of *C. reinhardtii* that had non-functional CVDE displayed constant qZ response even in the absence of a pH gradient across the thylakoid membrane (Holub *et al*., 2007).

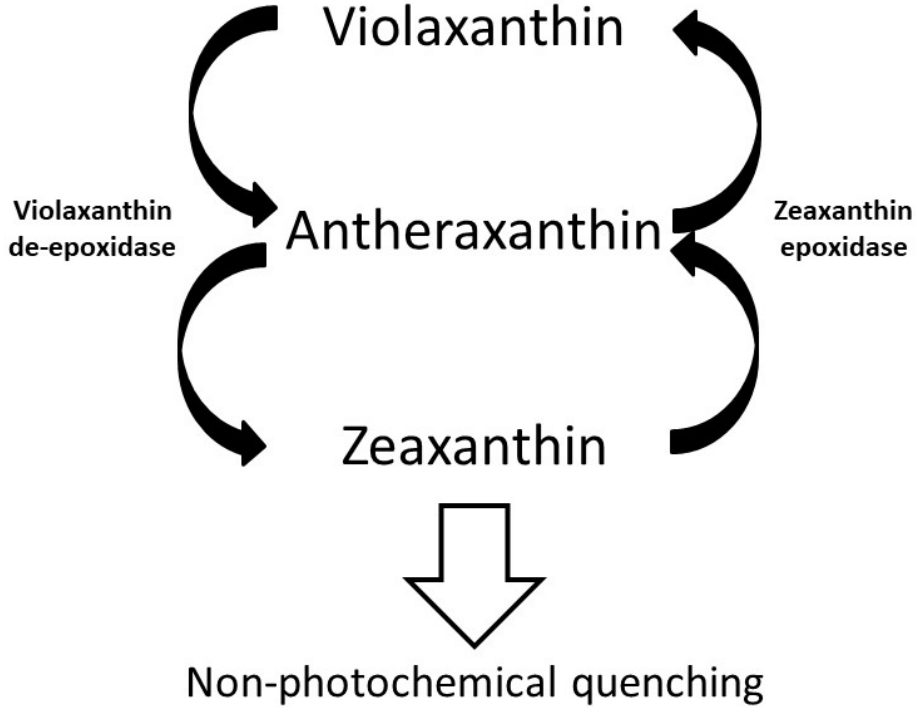


Figure 1.6: The xanthophyll cycle in C. reinhardtii. Violaxanthin is converted to zeaxanthin via an antheraxanthin intermediate by violaxanthin de-epoxidase. Zeaxanthin can then perform photoprotection through NPQ. Zeaxanthin can then be converted back to violaxanthin via anterhaxanthin by the enzyme zeaxanthin epoxidase.

The final short-term NPQ response is through state transitions (qT). Under high light stress, it is beneficial to have more light harvesting complexes (LHC) surrounding PSI than PSII, as PSI is quicker to quench excited chlorophyll and is less susceptible to photodamage (Papageorgiou, 2014). qT refers to the adjustment of the distribution of LHCs between PSII and PSI, and is regulated in *Arabidopsis thaliana* by the opposing functions of STN7 kinase and phosphorylase TAP38 (Allen, 1992; Pribil *et al*., 2010). In *C. reinhardtii,* the kinase involved in qT is STT7, and no phosphorylase has been identified to date (Depège *et al.*, 2003; Dumas *et al.*, 2017). In state 1 (no high light stress, or when light wavelengths are enriched for PSI), LHCI is associated mainly with PSI and LHCII is mainly associated with PSII. An over-reduced plastoquinone pool from high light exposure leads to binding of plastoquinol at the Q_0 site of cytochrome b_6f, which in turn activates STT7 (Zito *et al.*, 1999; Lemeille *et al*., 2009). STT7 then phosphorylates LHCII, which dissociates from PSII and either associates with PSI or forms aggregates in the thylakoid membrane (state 2; Ünlü *et al*., 2014). This

process can then be reversed by a phosphorylase. The aggregation of LHCII is believed to contribute to reduced high light stress by removing it from either photosystem, and has also been suggested to contribute to qE quenching (Ruban and Johnson, 2009; Betterle *et al.*, 2009; Ünlü *et al.*, 2014). The state transition shifts have been shown to be greater in *C. reinhardtii* than in higher plants in terms of percentage of the LHCII antennae that dissociate from PSII (Vallon *et al.*, 1986; Allen, 1992). Therefore qT is considered a short term response in *C. reinhardtii*, whereas it is considered longer term in higher plants. The balance between qE and qT contribution to NPQ has also been linked to photoautotrophic and mixotrophic growth, with the former leading to a greater effect of qE and the latter a greater effect of qT (Finazzi *et al.*, 2006; Iwai *et al.*, 2007).

A combination of photochemical and non-photochemical quenching therefore contributes to the photoprotection of *C. reinhardtii*, particularly under fluctuating light conditions.

1.2.2.3. Long term strategies of photoprotection

Long term responses in *C. reinhardtii* to changing light conditions allow for survival upon extended periods of high light intensity, for example due to seasonal changes. Most of these changes are due to gene expression changes, which have been shown to be in response to increased singlet oxygen presence in the chloroplasts and an over-reduced plastoquinone pool (Teramoto *et al*., 2002; Ledford *et al*., 2007). In fact, it has been shown that exposure to a small amount of singlet oxygen will 'prime' a culture and allow it to survive higher concentrations of singlet oxygen after a few hours (Ledford *et al*., 2007). The retrograde signalling pathway that results in nuclear gene expression responding to singlet oxygen generated in the chloroplast by high light stress remains mostly unknown (Erickson *et al*., 2015). A singlet oxygen response regulator has been identified named SAK1, which has been shown to be essential for the induction of many genes by singlet oxygen (Wakao *et al*., 2014). SAK1 has been found to be a cytosolic protein phosphorylated upon singlet oxygen exposure, and is therefore likely to be a part of a kinase retrograde signalling pathway connecting the chloroplast to nuclear gene expression. A transcription factor called SOR1 had also been identified using a mutant line where the protein was overactive and provided increased protection against singlet oxygen stress (Fischer *et al*., 2012). The genes induced by this mutant line matched those induced by reactive electrophilic species (RES), a group of highly reactive chemicals that act to affect gene expression in response to a variety of stresses, including high light stress, as ROS have been observed to generate RES (Havaux *et al.*, 2005; Farmer *et al*., 2007). Therefore it is thought that there is overlap between singlet oxygen and RES signalling pathways. However, the

genes regulated by SOR1 and SAK1 were found not to overlap completely (Wakao *et al*., 2014), suggesting there are multiple retrograde pathways responding to singlet oxygen and high light stress in *C. reinhardtii*. This is supported by the observation that treating *C. reinhardtii* cultures with RES only partly increased subsequent resistance to singlet oxygen exposure, unlike the 'priming' experiments using singlet oxygen pre-treatment (Ledford *et al*., 2007; Fischer *et al*., 2012). Methylene blue sensitivity (MBS) protein has also been implicated in singlet oxygen signalling (Shao *et al*., 2007). Despite being constitutively expressed, MBS instead aggregates along with untranslated mRNA in response to singlet oxygen (Shao *et al*., 2013), and could thus perform a role in amplifying the signalling response.

RNA sequencing has been performed on *C. reinhardtii* after exposure to singlet oxygen stress (Wakao *et al.*, 2014). Transcripts associated with ABC-transporters related to drug resistance were found to be upregulated, which was thought to be linked to removing photosensitising molecules from the cell in an aquatic environment. Glutathione peroxidase (GPXH) has also been found to be greatly upregulated upon singlet oxygen exposure, and has therefore often been used as a positive control for ROS stress in *C. reinhardtii* (Fischer *et al*., 2006). GPXH functions to reduce peroxides, in particular hydrogen peroxide, and is thus useful in the response to high light stress and ROS control (Fischer *et al*., 2009). LHC antenna protein expression is also regulated in the high light response, as a decreased antenna size around the photosystems will result in decreased photodamage under high light stress (Durnford *et al.*, 2003; McKim *et al.*, 2006). Interestingly, regulation of LHC protein expression only occurs over the first few hours of high light exposure in *C. reinhardtii*, with LHC levels recovering around 6 hours after irradiation (Durnford *et al*. 2003).

A strong link has also been observed between light intensity and regulation of carbon concentration mechanisms in *C. reinhardtii* (Marcus *et al.*, 1986; Wang *et al.*, 2015). CCM acts to increase CO_2 concentration around rubisco, thus improving the efficiency of carbon reduction and assimilation into the Calvin cycle (Moroney *et al.*, 1985; Raven *et al.*, 2008). Therefore increasing CCM when the cell is exposed to light would allow for more CO_2 to act as a sink for reducing equivalents produced by the PETC, thus alleviating a back-up in linear electron flow that could lead to ROS production. Equally, it could be that increasing light intensity and therefore reducing equivalents from the PETC leads to faster CO_2 uptake in the Calvin cycle, which leads to a lower CO_2 concentration around rubisco, and this reduction in CO_2 concentration initiates an increase in CCM. It has been shown that CCM components are upregulated in the dark to light transition during the diurnal cycle (Mitchell *et al.*, 2014; Tirumani *et al.*, 2014), further demonstrating the link between CCM and photosynthesis in *C. reinhardtii*. CCM will be reviewed in more detail in chapter 7.

Another long term photoprotective element is the increased production of antioxidants. A combination of increased enzyme activity and expression results in a large pool of both carotenoids and tocopherols within a few hours of high light exposure (Baroli *et al.*, 2003; Bonente *et al*., 2012; Li *et al.*, 2012). These antioxidants continue to quench singlet oxygen and excited chlorophyll to counter damage the former will cause.

A long term form of NPQ is known as photoinhibition (qI), and involves the repair of PSII subunit D1. This subunit is particularly susceptible to high light damage as it contains the reaction centre and special pair of chlorophyll. The repair system removes damaged D1 from PSII and destroys it via FtsH protease (Aro *et al.*, 1993; Malnoë *et al.*, 2014). However, high light has been shown to overwhelm the repair system, which reduces the flow of electrons into the PETC (Takahashi *et al*., 2005; Murata *et al*., 2012). Despite seeming detrimental to decrease PSII efficiency under stress, this is thought to protect PSI from further damage from increased election flow, and therefore decrease the effect of high light stress (Tikkanen *et al*., 2014).

1.3. Cytochrome c_{6A} could be involved in photoprotection in *C. reinhardtii*

1.3.1. Discovery of cytochrome c_{6A}

Cytochrome c_6 is a haemoprotein that serves in cyanobacteria and *C. reinhardtii* as a functional substitute to plastocyanin in copper-depleted conditions (Wood, 1978; figure 1.2). Prior to the 21st century, cytochrome c_6 was not thought to exist in higher plants, with plastocyanin completely replacing it functionally (Kerfeld *et al.*, 1998). However, in 2002 two laboratories independently found evidence for a higher plant homologue of cytochrome c_6 (later named cytochrome c_{6A}). Gupta *et al*. (2002) discovered a cytochrome c_6 homologue during a yeast-2-hybrid screen performed in *Arabidopsis thaliana*. The novel cytochrome bound to the plastid immunophilin FKBP13. The study found that cytochrome c_{6A} knockout mutant lines in *A. thaliana* showed no clear phenotypes, but RNAi knockdown of plastocyanin in a mutant lacking cytochrome c_{6A} was lethal. This led Gupta *et al*. (2002) to suggest that cytochrome c_{6A} and plastocyanin were functionally the same, and that the former performed a similar role to cytochrome c_6 in cyanobacteria and unicellular eukaryotic algae.

The other lab to report evidence of a cytochrome c_6 homologue in higher plants found it through searching genomic and expression sequence tag (EST) databases of many land plants, including *A. thaliana* (Wastl *et al*., 2002). Sequences were found that could encode a protein that was similar in sequence to cytochrome c_6 and included a haem-binding motif (CXXCH) and a methionine residue that could potentially ligate to the iron in a haem prosthetic group. However, Wastl *et al*. suggested

that cytochrome c_{6A} may not perform a similar function to cytochrome c_6 as Gupta *et al.* (2002) had hypothesised, as the calculated surface charge of this novel cytochrome differed significantly from cytochrome c_6 (Figure 1.7). One aspect of cytochrome c_{6A} that both studies highlighted was an insertion of 12 amino acids containing two cysteine residues into a loop region. This was named the loop insertion peptide (LIP; Howe *et al*., 2005). It was predicted that these two cysteines formed a disulphide bridge, which was confirmed in later studies using x-ray crystallography (figure 1.7; Marcaida *et al.*, 2006). The crystal structures of cytochrome c_{6A} showed it consisted of 4 alpha helices joined by turns, and also showed the presence of the haem prosthetic group, coordinated between the methionine and histidine residues of the motifs mentioned before. The crystal structure showed little difference between cytochrome c_{6A} and cytochrome c_6, except for the extra LIP in the former. There was also little difference found between cytochrome c_{6A} crystal structures where the haem was oxidised or reduced (Marcaida *et al*., 2006).

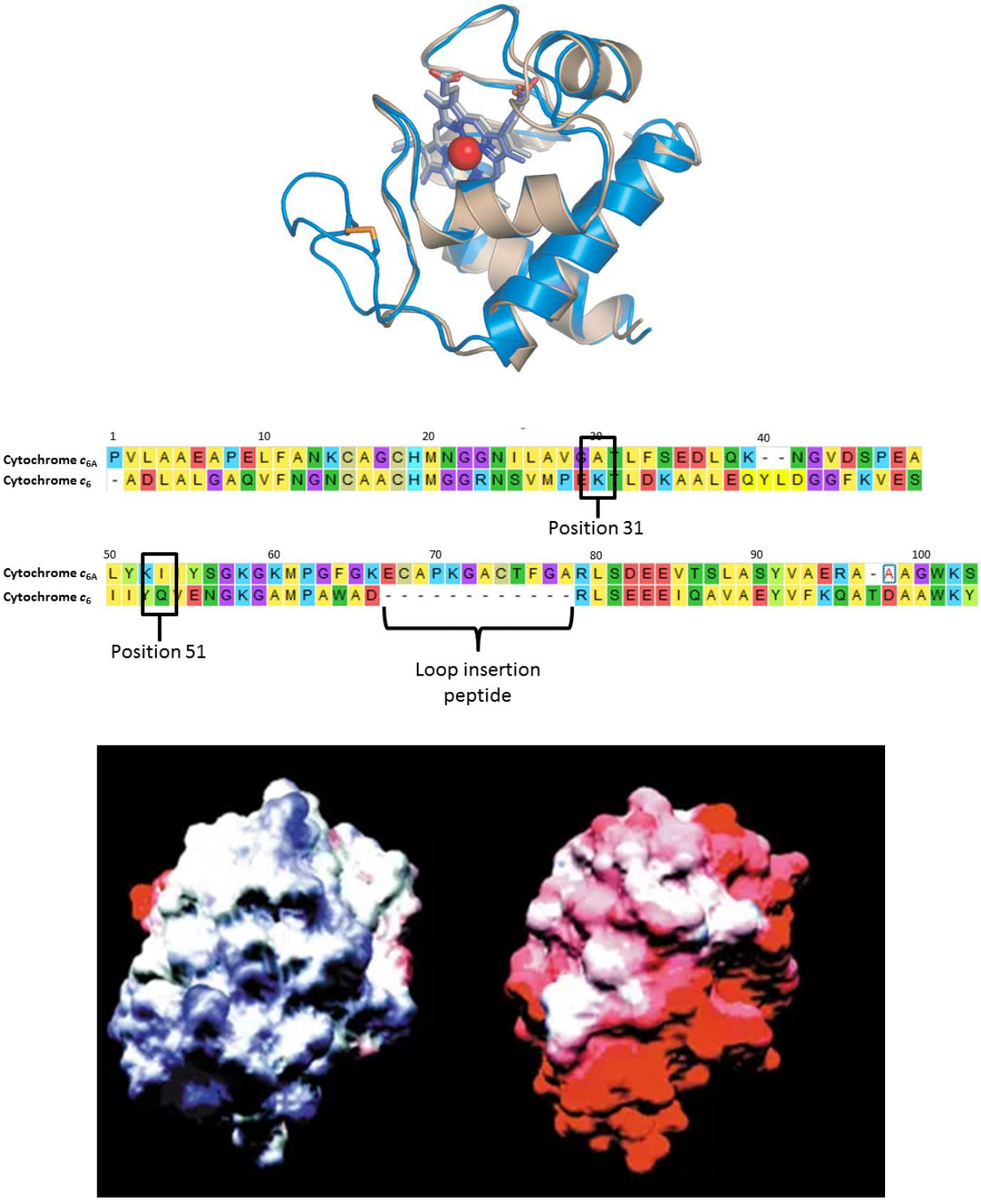


Figure 1.7: Crystal structure (upper), peptide sequence (middle) and calculated surface electrostatic potential distribution (lower) of cytochrome c_{6A} and cytochrome c_6. For the crystal structure, Arabidopsis thaliana cytochrome c_{6A} (blue) is superimposed onto the crystal structure of Monoraphidium braunii cytochrome c_6 (beige). The haem group is coloured dark blue, with a red iron ion, and the loop insertion peptide has been highlighted with a yellow disulphide bridge. For the alignment, cytochrome c_{6A} and cytochrome c_6 from Chlamydomonas reinhardtii have been aligned

using ClustalW in MEGA7. Positions 31, 51 and the loop insertion peptide in cytochrome c_{6A} have been highlighted. Positively charged residues have been coloured blue, hydrophobic residues yellow, negatively charged residues red, polar residues green, cysteines dark yellow, and glycines purple. For the surface charge distribution, Arabidopsis thaliana cytochrome c_{6A} is shown on the left, and Monoraphidium cytochrome c_6 on the right. Negative and positive charge are shown by red and blue surfaces respectively. Adapted from Molina-Heredia et al., 2003 and Marcaida et al., 2006.

1.3.2. Evidence that cytochrome c_{6A} is functionally different from cytochrome c_6

As more studies were performed on cytochrome c_{6A}, it became increasingly clear that it cannot perform the same function as cytochrome c_6. First, cytochrome c_{6A} sequences with the LIP were found in *C. reinhardtii* (Wastl *et al.*, 2004a), an organism that already has a cytochrome c_6. The fact that both cytochromes exist in some organisms suggests that they may perform different functions. Weigel *et al*. (2003) found that knocking out both alleles of plastocyanin in *A. thaliana* proved lethal unless the plants were grown on media supplemented with sucrose. This phenotype was not complemented by overexpressing cytochrome c_{6A}, thus countering the observations of Gupta *et al*. (2002). A reason for this disagreement could be that the earlier study used knockdown lines of plastocyanin, and that therefore there was enough plastocyanin to allow (stunted) growth of plants. This protein level was too low for the sensitivity of the western blot performed. However, these experiments suggest that the lack of cytochrome c_{6A} is lethal under conditions resulting from a low level of plastocyanin.

Based on detailed modelling work, cytochrome c_{6A} was also predicted to have a different surface electrostatic-potential distribution from cytochrome c_6 and plastocyanin (Figure 1.7; Molina-Heredia *et al.*, 2003). This supported the prediction from Wastl *et al.* (2002), and strongly suggested that cytochrome c_{6A} has different binding partners and kinetics than cytochrome c_6 and plastocyanin. Cytochrome c_{6A} was also found to have a redox midpoint potential over 200 mV lower than cytochrome c_6 (Worral *et al*., 2007). This means that the novel cytochrome would not be able to oxidise the cytochrome *f* component of cytochrome b_6f, but could oxidise plastoquinol (figure 1.8A). Cytochrome c_{6A} could still theoretically reduce PSI, and this interaction has been tested *in vitro* (Molina-Heredia *et al.*, 2003; Wastl *et al.*, 2004b). It was found that cytochrome c_{6A} was much slower to reduce PSI than cytochrome c_6 or plastocyanin, even under conditions that would reduce the effect of the electrostatic differences between the proteins. Cytochrome c_{6A} has also been shown to be able to reduce plastocyanin *in vitro* (Marcaida *et al*., 2006). The lower redox midpoint potential of

cytochrome c_{6A} has been associated with two specific amino acid residues: a valine and an alanine at positions 52 and 31 respectively in *A. thaliana* (figure 1.7; Worrall *et al*. 2007). The V52 residue was found to contribute more than A31 to the reduction in midpoint potential, with a V52Q mutant form of cytochrome c_{6A} having a midpoint potential ~100 mV higher than the native form. In cytochrome c_6 Q52 is likely to form a hydrogen bond with the haem propionate side chain via a water molecule, which valine would be unable to form. Replacing residue A31 on the other hand raised the midpoint potential by only ~30 mV. The LIP was found not to contribute significantly to the redox midpoint potential, as deleting it or replacing the cysteines with serines did not bring the p*I* of the protein close to that of plastocyanin (Wastl *et al*., 2004b). These observations collectively indicate that cytochrome c_{6A} plays a different role from cytochrome c_6.

1.3.3. Previously suggested functions for cytochrome c_{6A}

If cytochrome c_{6A} is not a functional substitute for cytochrome c_6, the question of its true function remains. Previously it has been suggested that cytochrome c_{6A} may function in apoptosis signalling (Weigel *et al.*, 2003). Cytochrome *c* has been shown to be an amplifying signal for apoptosis in eukaryotes, being released from mitochondria and interacting with caspases (Li *et al.*, 1997). However, proteomic analysis of *A. thaliana* undergoing apoptosis showed no change in expression of cytochrome c_{6A} (Swidzinski *et al*., 2004; Howe *et al*., 2006).

Another suggestion was that the cysteines in the LIP could assist in disulphide bridge formation in other thylakoid proteins (Schlarb-Ridley *et al.*, 2006). The disulphide bridge of cytochrome c_{6A} would oxidise two cysteines in the target protein, leaving two thiol groups. One thiol could then reduce plastocyanin, leaving a disulphide anion radical. The electron from this radical would then transfer to the haem group, resetting the disulphide bridge in the LIP. The haem would then reduce another plastocyanin molecule, allowing another cycle to begin. This partnership may also suggest why a knockdown of plastocyanin and cytochrome c_{6A} in *A. thaliana* was found to be lethal by Gupta *et al.* (2002). However, Worrall *et al*. (2008) found through analysis of the angles of the disulphide bond in cytochrome c_{6A} that the LIP had more of a structural than catalytic role, using the observations of Schmidt *et al*. (2006) on the relationship between disulphide bridge angle and function. A structural role for the LIP was supported by the observation that replacing the cysteines of the LIP with serines proved significantly detrimental to cytochrome c_{6A} structural stability (Mason *et al*., 2012). Therefore a catalytic role involving thiol redox has been deemed unlikely, but cannot be ruled out.

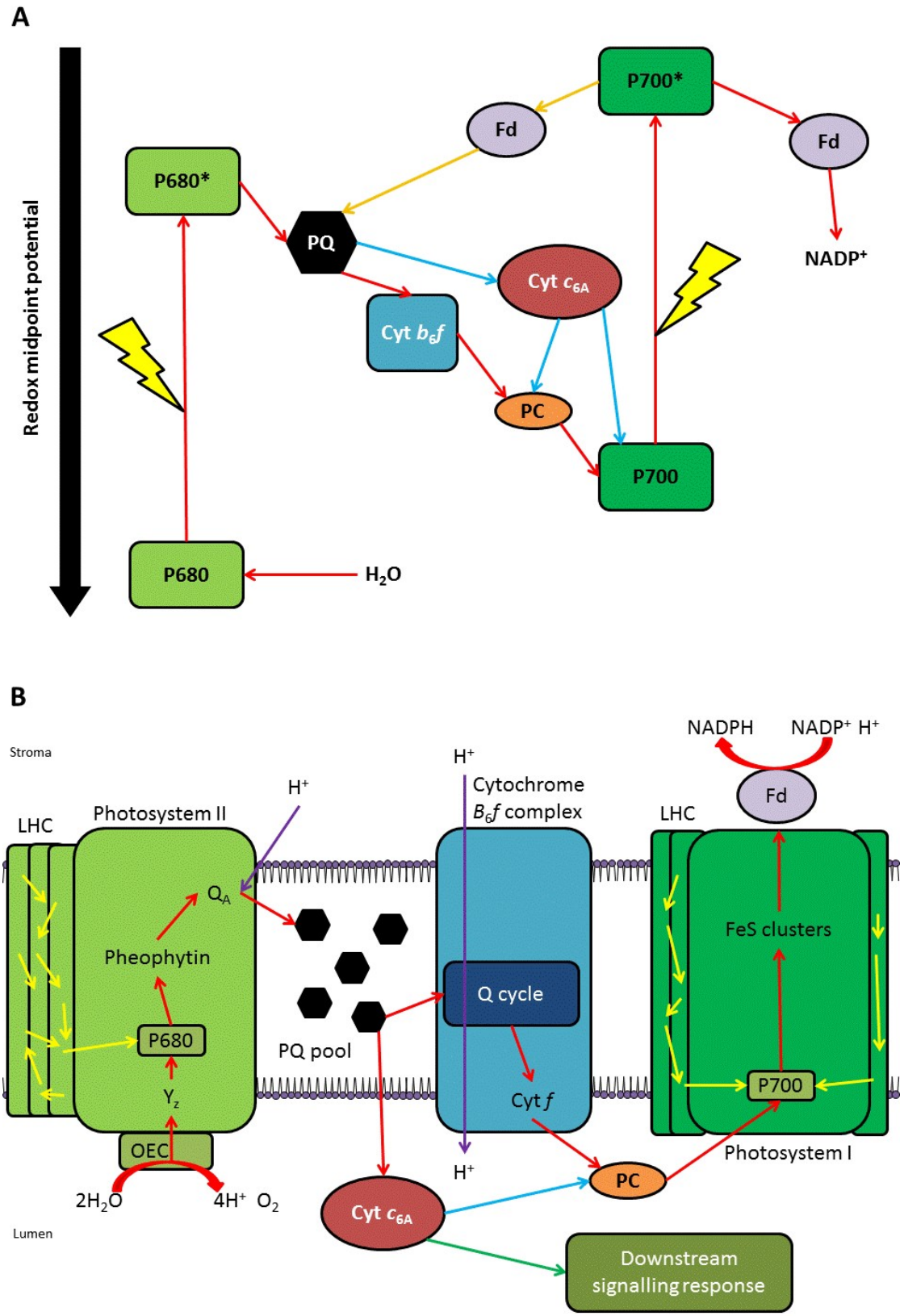


Figure 1.8: How cytochrome c_{6A} may fit into the photosynthetic electron transfer chain. A) Diagram of the Z-scheme of photosynthesis, with red arrows representing linear electron flow, orange arrows representing cyclic electron flow, and blue arrows representing potential redox targets of cytochrome

c_{6A}. Lightning symbols represent where electrons are excited at photosystem reaction centres by light energy. B) Diagram of the thylakoid membrane and photosynthetic machinery, with red arrows representing linear electron flow, purple arrows representing proton movements, blue arrows representing potential electron flow in the 'safety valve' hypothesis of cytochrome c_{6A} function, and the green arrow representing the signalling hypothesis of cytochrome c_{6A}, in which a downstream signalling molecule is reduced to activate a high light stress response. Abbreviations in the diagrams are: Cyt cytochrome; PQ plastoquinone; PC plastocyanin; Fd ferredoxin; OEC oxygen evolving complex.

1.3.4. Current hypotheses for cytochrome c_{6A} function

There are currently two main hypotheses for cytochrome c_{6A} function that this study wishes to investigate (figure 1.8B). The first is that cytochrome c_{6A} could act as a 'safety valve', in a similar way to the other alternative electron pathways discussed in section 1.2.2.2. Although it has a lower redox midpoint potential than cytochrome c_6 it could nevertheless oxidise plastoquinol in theory, and transfer these electrons to either plastocyanin or PSI. It has already been shown that cytochrome c_{6A} can reduce both plastocyanin and PSI (Molina-Heredia *et al.*, 2003; Marcaida *et al.*, 2006), and this would bypass cytochrome b_6f, which has been shown to be a bottleneck in the PETC due to the Q cycle (Papageorgiou, 2007). Therefore under high light conditions, cytochrome c_{6A} could be contributing to the extra passage of electrons through the PETC to increase photochemical quenching and alleviate ROS production at PSII. This hypothesis has the assumption that cytochrome c_{6A}, which has been shown to reside in the thylakoid lumen (Nimmo, 2011); can interact with plastoquinol, which exists in the thylakoid membrane. However, other studies have shown that the quinol headgroup of ubiquinol and plastoquinol can push through the membrane to exchange electrons with respiratory complex I and chloroplast sensor kinase respectively (Baradaran *et al.*, 2013; Puthiyaveetil *et al.*, 2013). Therefore an interaction between plastoquinol and cytochrome c_{6A} is not impossible.

The second main hypothesis for cytochrome c_{6A} function is that it could function in the signalling network in response to high light. There is evidence, for example, that the redox state of the plastoquinone pool can cause changes in gene expression (Teramoto *et al.*, 2002). Cytochrome c_{6A} could function in removing electrons from an over-reduced plastoquinone pool, which would in turn activate a retrograde signalling pathway inducing nuclear expression change. Redox-regulated thiol reactions occurring in the thylakoid lumen have also been shown to contribute to photoprotection

through xanthophyll cycle and PSII turnover control (Hall *et al.*, 2010), therefore suggesting that cytochrome c_{6A} thiol redox chemistry could actually be involved. One argument against this hypothesis is the observation from X-ray crystallography that cytochrome c_{6A} does not significantly alter in structure upon haem reduction (Marcaida *et al*., 2006). This means that cytochrome c_{6A} would not necessarily change interaction partners upon reduction. However, cytochrome c_{6A} may just reduce another protein in the pathway, which could subsequently activate through a change in structure. This crystallography and subsequent surface charge redox distribution was also performed on cytochrome c_{6A} at pH of around 7.5-8.5, which is higher than the pH of the thylakoid lumen, which can go down as far as 4.0 in high light stress (Jagendorf and Uribe, 1966; Krieger and Weiss, 1993). This could mean that at neutral pH, cytochrome c_{6A} is not active, but at lower pH the structure, midpoint potential or surface charge distribution could be altered in a way that allows a pH-induced response. Another observation that lends support to the signalling hypothesis is that cytochrome c_{6A} has a relatively low abundance in the lumen (Nimmo, 2011). Therefore unless the expression levels of cytochrome c_{6A} change significantly under stress conditions, a low expressed protein would fit a signalling role more than a safety valve role as there is opportunity for amplification.

1.3.5. Discovery of cytochromes c_{6B} and c_{6C} in cyanobacteria

A low-redox midpoint potential cytochrome c_6 was discovered in the cyanobacterium *Synechochoccus sp.* PCC 7002 through Southern blot screening using the native cytochrome c_6 sequence (Nomura, 2001). This cytochrome (named PETJ2) was found to have a redox midpoint potential of +155 mV, which is closer to that of cytochrome c_{6A} than to cytochrome c_6 (Bialek *et al.*, 2008). However, sequence analysis of the coding sequence showed there was not an equivalent of the LIP in PETJ2. Homologues of PETJ2 were also identified in a wide range of cyanobacteria. A phylogenetic analysis of PETJ2 peptide sequences across cyanobacterial species suggested that there were two distinct groups of PETJ2, one that had a more recent common ancestor with cytochrome c_{6A}, and one with a more recent common ancestor with cytochrome c_6 (Bialek *et al.*, 2008). These groups of proteins were named cytochrome c_{6B} and c_{6C} respectively. The discovery and potential function of cytochromes c_{6B} and c_{6C} will be reviewed in more detail in chapter 3.

1.4. Overall aims of the study

The aim of this study is to investigate cytochrome c_{6A} function, and specifically test whether either the safety valve hypothesis, the signalling hypothesis, both or neither hypotheses are true. Firstly, the ancestry and phylogeny of the cytochrome c_6 family are analysed bioinformatically, to gain a greater understanding of the evolutionary history and distribution of cytochrome c_{6A}, c_{6B} and c_{6C} amongst photosynthetic organisms.

Cytochrome c_{6A} knock out mutants in *C. reinhardtii* have been created using a CRISPR/Cpf1 system (Ferenczi *et al.*, 2017). This study aims to characterise these mutants by comparing their growth under stress conditions, and comparing their photosynthetic efficiency with that of their background strain cw15. The structure of cytochrome c_{6A} at different pH values is monitored using circular dichroism to see if there are any large-scale changes in cytochrome c_{6A} structure with changes in pH of the thylakoid lumen. Finally, the transcriptome response of *C. reinhardtii* to high light stress is also tested both in a cytochrome c_{6A} knockout strain and cw15 through RNA sequencing.

2. Materials and Methods

2.1. Algal sources

Chlamydomonas reinhardtii background strain cw15, a cell wall deficient mutant (Davies and Plaskitt, 1971), was sourced from Saul Purton (University College London). Cw15 was the background strain used for the c_{6A} knockout and knockdown strains.

Cytochrome c_{6A} knockout strains Uex-55, Uex-78 and Uex-87 were provided by Aron Ferenczi (University of Edinburgh) and were generated using a CRISPR/cpf1-mediated DNA editing system (Ferenczi *et al.*, 2017).

2.2. Bacterial sources

Escherichia coli strains were all sourced from laboratory stocks. Background strains DH5α and GM119 contain mutations for increased transformation efficiency (Hanahan, 1983 and Arraj *et al.,* 1983 respectively).

2.3. Plasmids used

Expression plasmids pATc6A and pEC86 were sourced from Adrian Barbrook (University of Cambridge), created as described in Wastl *et al.,* 2005 and Schulz *et al.*, 1999 respectively.

2.4. Antibodies used

The antibodies used in this study are summarised in table 2.1.

Antibody used	Source
Rabbit anti-cytochrome c_{6A} (*A. thaliana*) from sera.	From laboratory stock (originally made by Covalab UK).
Goat anti-rabbit (conjugated alkaline phosphatase).	AbCam (product code ab6722).

Table 2.1: List of antibodies and their sources used in this study.

2.5. List of suppliers providing reagents

The following suppliers provided chemical and biological reagents for the study:

Abcam

BioRad

Biotium

Duchefa

Eppendorf

Expedeon

Fisons

GE Healthcare

Greiner Bio-one

Honeywell

Invitrogen

Merck

New England Biolabs

Novogene

Qiagen

Promega

Sarstedt

Sartorius Stedim

Sigma Aldritch

Starlab

Thermo Fischer Scientific

2.6. Methods in *Chlamydomonas reinhardtii*

2.6.1. Growth of *C. reinhardtii*

2.6.1.1. TP and TAP recipes

Cultures of *C. reinhardtii* were grown in either Tris Phosphate (TP) or Tris Acetate Phosphate (TAP) medium, the recipes of which are found in table 2.2. All media were autoclaved prior to use.

Component	Amount
Beijerinck salts (40x; table 2.3)	25 mL
Hutner's Trace Elements (table 2.4)	1 mL
1 M KPO_4, pH 7 250 mL 1 M K_2HPO_4 170 mL 1 M KH_2PO_4	1 mL
Tris Base	2.42 g
dH_2O	Up to 1 L
Adjust to pH 7 with HCl (TP) or glacial acetic acid (TAP)	

Table 2.2: Recipe for TP and TAP media. To make TP or TAP the pH was adjusted with HCl or acetate respectively.

Component	Amount
NH_4Cl	16 g
$CaCl_2.2H_2O$	2 g
$MgSO_4.7H_2O$	4 g
dH_2O	Up to 1 L

Table 2.3: Recipe for Beijerinck salts.

Component	Amount
H_3BO_3	11.40 g
$ZnSO_4.7H_2O$	22.00 g
$MnCl_2.4H_2O$	5.06 g
$FeSO_4.7H_2O$	4.99 g
$CoCl_2.6H_2O$	1.61 g
$CuSO_4.5H_2O$	1.57 g
$(NH_4)_6Mo_7O_{24}$	1.10 g
dH_2O	Up to 1 L

Table 2.4: Recipe for Hutner's Trace Elements.

2.6.1.2. Growth on solid media

Solid media plates were made through autoclaving TP or TAP media with 2% w/v Plant Agar (Duchefa). Solid media stocks for all strains were kept unless otherwise stated in a Gallenkamp cooled static incubator at 30 °C under 40 µE m^{-2} s^{-2} white light.

Spot tests were performed by pipetting 20 µL liquid culture at OD_{750} ~ 0.5 onto solid medium with regular spacing, then keeping under the conditions described above unless otherwise stated.

2.6.1.3. Growth in liquid media

Single colonies were taken from solid media plates and suspended in 10 mL of either TP or TAP medium. After growth to mid-log phase cultures were then transferred to 50 mL TP or TAP media. Any glassware used for growing cultures was submerged in 1 M HCl overnight and cleaned with distilled water before autoclaving.

Unless otherwise stated liquid cultures were grown in a New Brunswick Innova G230 refrigerated incubator at 30 °C under 40 µE m^{-2} s^{-1} white light with 200 rpm shaking. Optical density measurements were made using the Helios δ spectrophotometer for cultures grown in glass flasks, and using the Shimadzu UV-1800 spectrophotometer for cultures grown in plastic Nunc flasks. Cell count measurements were made by adding 10 µL of culture containing 10 % Lugol solution to a haemocytometer.

Liquid cultures grown in the Algem (Algenuity) were grown in 500 mL TP medium, with 120 rpm shaking and 30 °C under 40 $\mu Em^{-2} s^{-1}$ white light.

2.6.2. Cryogenic storage of *C. reinhardtii*

C. reinhardtii cells grown on TAP plates were restreaked twice at two-day intervals. Cells were then resuspended in 5mL of TAP-sucrose medium (80 % TAP and 20 % 40 mM sucrose) to a concentration of $1x10^7$ cells/mL. In dim light the suspension was then diluted again to $2x10^6$ cells/mL before adding to a tube of TAP-MeOH medium (90 % TAP, 10 % Methanol) in a ratio of 1:1. The tubes were then sealed, placed in ice, then cooled at 1 °C/min in a Mr Frosty Freezing Container (Nalgene) for 90 minutes. Frozen tubes were then transferred to liquid nitrogen and stored until needed.

To thaw and plate the cells, the tubes were removed from the liquid nitrogen and placed at 30 °C for 2 minutes with gentle shaking. Aliquots of 10 µL were then mixed with 2 mL TAP medium containing 0.5 % w/v agarose at 42 °C and plated on TAP plates. The plates were left in dim light for 1 hour at room temperature, sealed, and placed under standard conditions for solid media.

2.6.3. DNA extraction from *C. reinhardtii*

Samples of 10 mL were taken from liquid culture once grown to A_{750} of 0.3. Genomic DNA was extracted using the GeneJet Genomic DNA Purification Kit (Thermo), according to the Gram-Negative Bacteria Genomic DNA Purification Protocol. DNA concentration and 260/230 and 260/280 ratios were estimated using a Nanodrop 1000.

2.6.4. RNA extraction from *C. reinhardtii*

Suspensions of *C. reinhardtii* cells were pelleted by spinning at 1600 xg for 5 minutes in the Hettich Universal 32 centrifuge. The pellet was resuspended in 1 mL Trizol and snap-frozen in liquid nitrogen. The suspension was thawed and mixed 1:1 with ice cold ethanol. The suspension was then purified using the Direct-zol RNA miniprep kit (Zymo Research) according to manufacturer's instructions, with the optional DNA digestion step. RNA concentration and 260/230 and 260/280 ratios were estimated using a Nanodrop 1000.

2.6.5. Protein extraction from *C. reinhardtii*

Aliquots of 2 mL were taken from liquid culture and snap-frozen in liquid N_2 before storing at -80 ^{o}C. The samples were then thawed and vortexed for 1 minute with 150 mg 1 mm glass beads (acid washed). The sample was then pelleted by centrifugation at 14,000 xg and the pellet resuspended in ice cold acetone containing 10 % trichloroacetic acid and 0.07 % β-mercaptoethanol (v/v). After incubation at -20 ^{o}C overnight the samples were centrifuged for 30 minutes at 14,000 xg at 4 ^{o}C and washed twice with ice cold acetone. The pellet was dried and resuspended in 500 µL SDS protein buffer (recipe in table 2.5).

Component	Amount
Sodium dodecyl sulphate (SDS)	1.00 g
HEPES (pH 7.4)	5.95 g
dH_2O	Up to 500 mL

Table 2.5: Recipe for SDS protein buffer.

2.6.6. Chlorofluorometry

Chlorofluorometry readings were made using the CFImager from Technologica. The in-built protocol wizard was used to control the light profile, with fluorescence readings every 1.3 s. Liquid cultures of OD_{750} ~ 0.35 in TP media were added in 6 mL aliquots to a 6 well plate before reading. Details of the light conditions used for the fluorescence profiles are given in chapter 4.

2.7 Methods for *Escherichia coli*

2.7.1 Growth of *E. coli*

2.7.1.1 LB medium recipe

Lysogeny broth (LB) was made using the recipe shown in table 2.6. All media was autoclaved prior to use.

Component	Amount
Peptone	10 g
Yeast Extract	5 g
NaCl	5 g
dH_2O	Up to 1 L

Table 2.6: Recipe for LB media.

2.7.1.2. Growth on solid medium

E. coli cells were spread with a plastic spreader onto plates containing LB medium autoclaved with 2 % Bacto-Agar (BD, Thermo Fisher). Unless stated otherwise, plates were sealed with Parafilm and incubated in a New Brunswick Innova G230 refrigerated incubator at 37 °C on a static shelf.

2.7.1.3. Growth in liquid medium

Unless stated otherwise liquid cultures of *E. coli* were grown in LB medium in glass flasks in a New Brunswick Innova G230 refrigerated incubator at 37 °C shaking at 200 rpm.

2.7.2 Cryogenic storage of competent *E. coli*

A single colony from an LB agar plate was resuspended in 5 mL LB and incubated overnight. A culture was then seeded by transfer of 1 mL to 100 mL fresh pre-warmed LB and incubated until OD_{600} of ~ 0.5 was reached. The suspension was incubated on ice for 15 minutes before centrifugation for 10 minutes at 2250 xg at 4 °C. The pellet was resuspended in 10 mL TSS buffer (recipe in table 2.7) and divided into 100 µL aliquots. These aliquots were stored at -80 °C.

Component	Amount
Polyethylene glycol 8000	10.0g
$MgSO_4$	0.250 g
$MgCl_2$ (anhydrous)	0.203 g
LB powder (Formedium)	2.5g
H_2O	Up to 1 L

Table 2.7: Recipe for TSS buffer.

2.7.3. Transformation of *E. coli*

Competent cells were thawed on ice, before ~100 ng DNA in H_2O was added. After incubation for 20 minutes on ice the cells were incubated for 1.5 minutes at 42 °C followed by incubation on ice again for 2 minutes. The cells were then suspended in 1 mL sterile LB medium and incubated at 37 °C for 1 hour as described in section 2.7.1.3. An aliquot of 100 µL was then spread on solid LB agar plates containing an appropriate antibiotic and incubated overnight as described in 2.7.1.2. Antibiotics used in this study are ampicillin and chloramphenicol at working concentrations of 100 µg/mL and 25 µg/mL respectively.

2.7.4. Plasmid DNA extraction from *E. coli*

Colonies from transformation plates were picked and suspended in 5 mL LB medium containing either ampicillin, chloramphenicol or both and grown overnight as described in section 2.7.1.3. The culture was then spun at 1600 xg for 5 minutes in a Hettich Universal 32 centrifuge. Plasmid DNA was extracted from the pellet using an ULTRACLEAN plasmid miniprep kit (Mo Bio) according to manufacturer's instructions. DNA content, 260/230 and 260/280 ratios were estimated using a Nanodrop 1000.

2.7.5. Protein extraction from *E. coli*

2.7.5.1. Raw protein extraction

Competent GM119 cells were transformed with both pEC86 (containing a construct for constitutively expressing the haem maturation cassette in bacteria; Wastl *et al.,* 2005) and pATc6A (containing a construct for expressing cytochrome c_{6A} in *Escherichia coli*; Schulz *et al.*, 1999) and plated on LB agar medium containing both 100 µg/mL ampicillin and 25 µg/mL chloramphenicol as described in 2.7.3. All subsequent steps had the same concentrations of ampicillin and chloramphenicol in liquid culture. Colonies picked from the plate were grown overnight in 5 mL LB medium as described in 2.7.1.3. Cultures were transferred to 1.25 L of LB containing 0.5 µM $FeCl_3$. The culture was incubated at 30 °C for 78 hours with shaking at 200 rpm.

Cells were spun down for 20 minutes at 30,500 rpm at 18 °C in a Beckman centrifuge with the JA-14 rotor. The pellet was suspended in 50 mL sucrose buffer (recipe in table 2.8) and mixed on a magnetic stirring plate for 20 minutes at room temperature.

Component	Amount
Sucrose	340.00 g
Tris HCl pH 7.5	2.42 g
Na_2EDTA	0.29 g
H_2O	Up to 1 L

Table 2.8: Recipe for sucrose buffer.

The suspension was spun down for 20 minutes at 30,500 rpm at 18 °C in a Beckman centrifuge with the JA-14 rotor. The pellet was suspended in 50 mL osmotic shock buffer (recipe in table 2.9) and mixed on a magnetic stirring plate for 20 minutes at 4 °C. The suspension was then sonicated on ice using a Sonicator Ultrasound Processor XL (Heat Systems). The mixture was sonicated for 5 minutes in 30 second pulses of 15 % with 30 second breaks between each pulse. The sonicated mixture was spun down for 30 minutes at 30,500 rpm at 18 °C in a Beckman centrifuge with the JA-14 rotor. The supernatant was then filtered through a Minisart 0.22 µm filter (Sartorius Stedim) and stored at -80 °C.

Component	Amount
Tris HCl pH 7.5	0.121 g
$MgCl_2$ hexahydrate	0.095 g
H_2O	Up to 1 L

Table 2.9: Recipe for osmotic shock buffer.

2.7.5.2. Purification through FPLC

Filtered *E. coli* lysate was firstly loaded into a GE Healthcare 16/10 XL ion exchange column controlled by a GE Healthcare ÄKTA FPLC. Before loading, the column was washed with 20 mM Tris HCl buffer. Protein was then loaded onto the column before a gradient was run from the prior buffer to 10 mM Tris HCl, 200 mM NaCl at a flow rate of 1 mL per minute. The eluate was collected in 1.9 mL fractions. Absorption at 280, 413 and 555 nm was measured for all eluted fractions using the in-built spectrophotometer and a Hitachi U-3010, and those that absorbed at 550 nm were collated and concentrated to below 1 mL using an Amicon 5000mW concentrator column.

The concentrated fractions were loaded by injection onto a 26/600 superdex75 size exclusion column. Before loading the column was washed with 10 mM Tris HCl, pH7.5 and 200 mM NaCl, and the same buffer was eluted through the column at a flow rate of 1 mL per minute until 300 mL of buffer had passed through. Absorptions at the same 3 wavelengths as above were measured for the eluate, with fractions absorbing at 550 nm being collected for further analysis.

2.8. Methods with DNA

2.8.1. PCR and electrophoresis

2.8.1.1. PCR through Phusion

The reaction mixture for Phusion PCRs is given in table 2.10, and the temperature profile given in figure 2.1. Specific primer sequences can be found in supplementary table S1.

Reagent	Amount
DNA	X µL
5X Phusion GC buffer	10.0 µL
dNTPs (10 mM each)	1.0 µL
Forward primer (10 µM)	2.5 µL
Reverse primer (10 µM)	2.5 µL
DNA polymerase (2 units/ µL)	0.5 µL
DMSO	1.5 µL
H_2O	Up to 50 µL

Table 2.10: Reaction mixture for Phusion PCR. Appropriate volume of DNA solution was added to achieve 0.2 µg DNA per reaction.

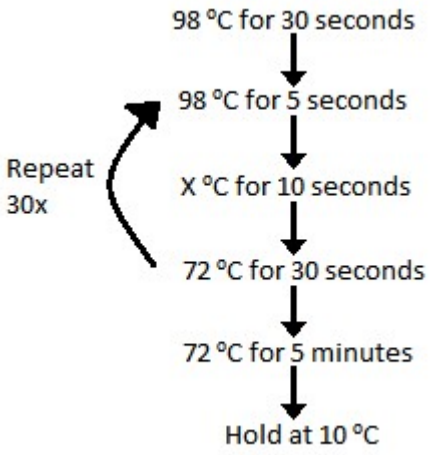


Figure 2.1: Temperature profile for Phusion PCR. Annealing temperatures (X) for specific primers given in supplementary information. The cycle number used was 30.

2.8.1.2. Agarose gel electrophoresis

Agarose gels were made by heating TBE buffer (see table 2.11) with 1 % agarose until melted. The molten mixture was poured into an appropriate mould with 1:50,000 diluted GelRed (Biotium). DNA samples were mixed with 6X DNA loading dye (Thermo Fisher Scientific) prior to loading. Unless otherwise stated gels were run at 150 mV for 30 minutes submerged in TBE buffer. Stained gels were visualised using an UGenius 3 UV imager (Syngene).

Component	Amount
Tris base	10.8 g
Boric acid	5.5 g
EDTA (pH 8.0)	20.0 mL of 0.5 M
dH_2O	Up to 1 L

Table 2.11: Recipe for TBE buffer.

2.8.2. Sequencing

2.8.2.1. Sanger sequencing

DNA solutions to be sequenced were prepared at a concentration of 100 ng/µL and sent with appropriate primers at a concentration of 10 µM to the DNA sequencing facility (Department of Biochemistry, University of Cambridge) for Sanger sequencing.

2.8.2.2. Whole Genomic sequencing

Extracted DNA sequences were analysed by Novogene using an Illumina machine running a paired-end 150 method. Resultant sequences were mapped to the *C. reinhardtii* genome from JGI (v5.5; Merchant *et al.*, 2007) and aligned using Geneious software.

2.8.3. Restriction digestion

Restriction digestions were set up as shown in table 2.12. Reaction mixtures were incubated at 37 ^{o}C for 1 hour in an Eppendorf Mastercycler Gradient.

Reagent	Amount
DNA	X µL
CutSmart buffer 10x (New England Biolabs)	5 µL
Restriction enzyme (each)	1 unit
H_2O	Up to 50 µL

Table 2.12: Restriction digestion mixture. DNA added (X) to a maximum of 1 µg.

2.9. Methods with RNA

2.9.1. Reverse transcription

Reverse transcription was performed on 1 µg total RNA using the superscript IV Reverse Transcription kit (Sigma Aldrich) according to manufacturer's instructions, with the optional inclusion of RNAseOUT (Sigma Aldrich). Negative control reactions (RT-) were performed with exclusion of the Reverse transcriptase.

2.9.2. RNA sequencing

RNA samples that demonstrated good purity and integrity through spectroscopy and electrophoresis respectively were aliquoted and sent to Novogene for sequencing using an Illimina NovaSeq and HiSeq platform. The transcriptome library (Eukaryotic Strand-Specific Transcriptome Library) was assembled according to the company's RNA-sequencing protocol.

The received data reads were trimmed according to quality of the read and to remove adapters using BBDuk (parameters shown in figure 2.3). Trimmed reads for each RNA sample were then mapped to the *C. reinhardtii* CDS annotated genome downloaded from JGI (v5.5; Merchant *et al.*, 2007). The mapping was performed on Geneious with the parameters outlined in figure 2.4.

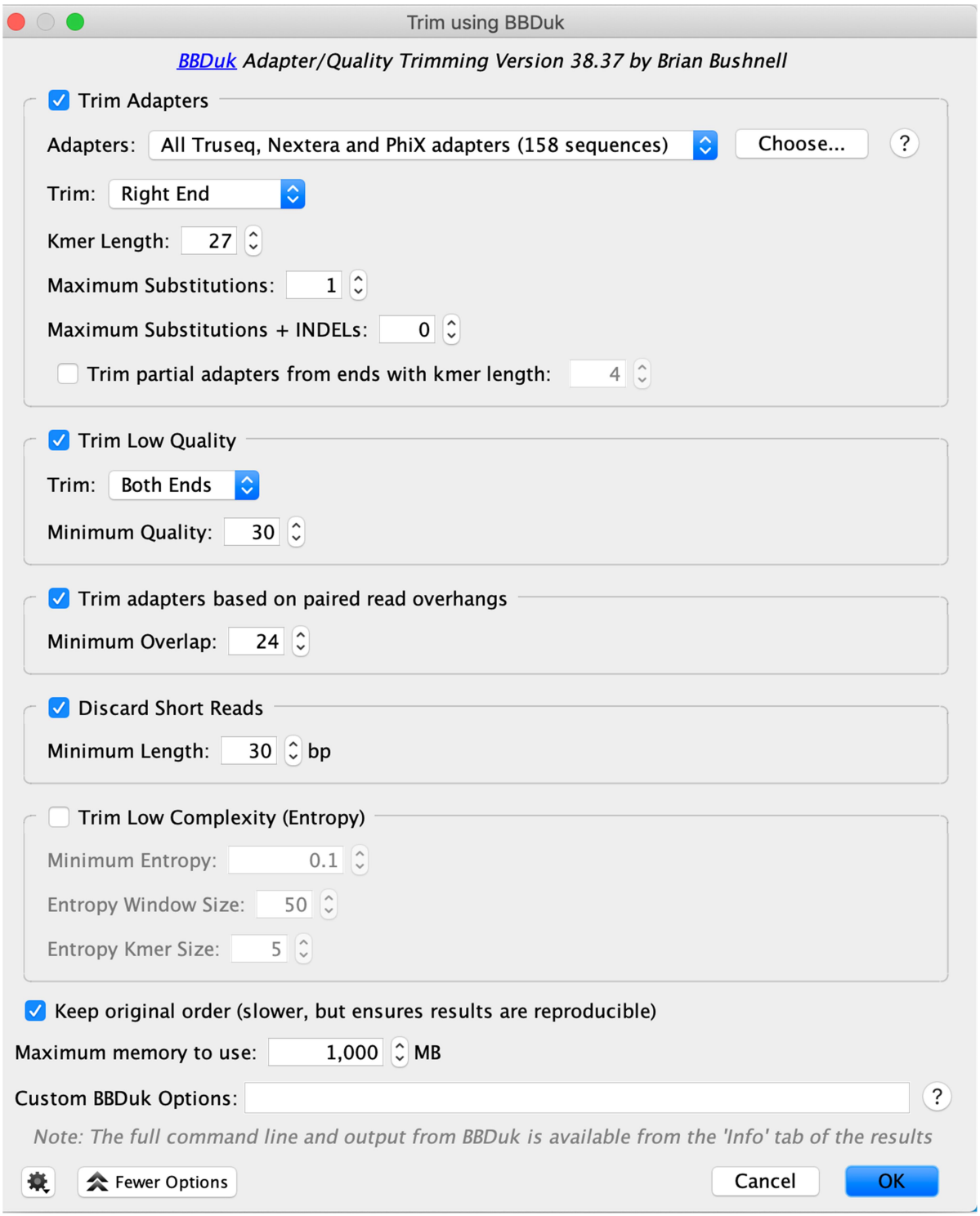


Figure 2.3: Geneious parameters used to trim the RNA reads prior to mapping to the C. reinhardtii genome.

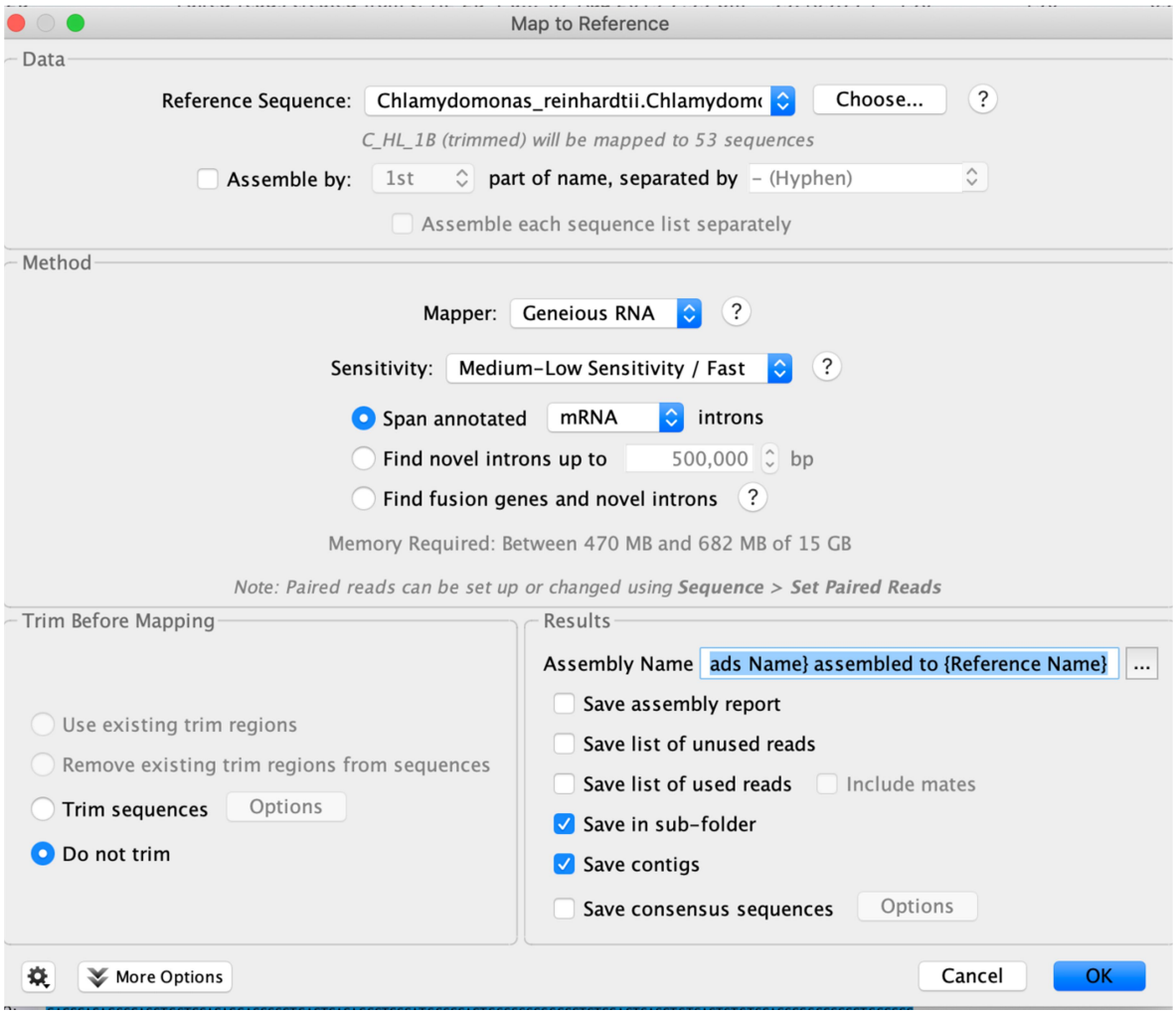


Figure 2.4: Geneious parameters used to Map the RNA sequencing reads to the C. reinhardtii genome downloaded from JGI (v5.5).

After mapping the reads, the expression levels were calculated through selecting “Annotate and Predict”, then from the drop-down menu “Calculate Expression Levels”. The annotation type was CDS. Once the expression levels for each sample were calculated, the data was compared through selecting “Annotate and Predict” followed by “Compare Expression Levels” in the drop-down menu. The DESeq2 method was then used to compare expression levels in each sample.

2.10. Methods with protein

2.10.1. Quantifying protein

Protein was quantified using the Pierce BCA Protein assay (Thermo) according to manufacturer's instructions. A Standard curve (figure 2.5) was prepared using Bovine Serum Albumin (BSA).

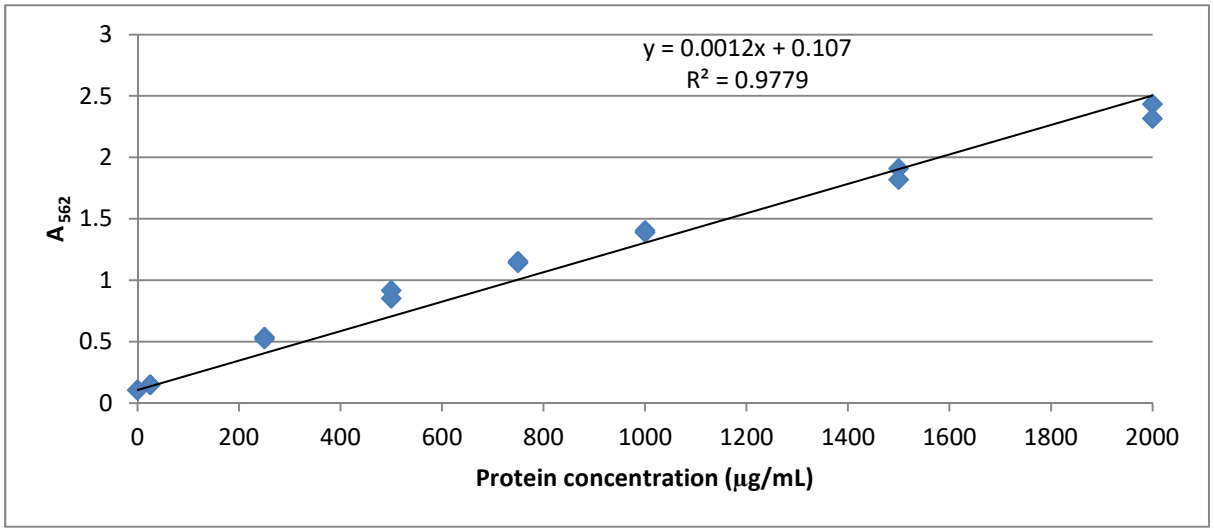


Figure 2.5: Standard curve relating protein concentration to A_{562} using the protocol from the Pierce BCA Protein assay. R^2 value and equation for the line of best fit were calculated by Microsoft Excel.

2.10.2. SDS-PAGE

Protein samples were diluted to a similar concentration in SDS protein buffer (table 2.5) and mixed 1:1 with 2X Protein loading buffer (recipe in table 2.13).

Component	Amount
Tris.HCl	381 mg
Glycerol	10 % (v/v)
B-Mercaptoethanol	5 % (v/v)
SDS	1 g
Methylene Blue	13 mg
H_2O	Up to 50 mL

Table 2.13: Recipe for 2X Protein loading buffer.

The mixture was incubated at 98 ^{o}C for 10 minutes before incubation on ice for 1 minute. Samples were then loaded into a Mini-Protean TGX Precast Gel (4-15 %), well size 50 µL. Gels were run in

SDS-PAGE running buffer (recipe in table 2.14) at 200 mV for 40 minutes, or until the dye front reached the bottom of the gel.

Component	Amount
Tris Base	15.1 g
Glycine	72.0 g
SDS	5.0 g
dH_2O	Up to 1 L

Table 2.14: Recipe for 5X SDS-PAGE running buffer.

SDS-PAGE gels were stained using Coomassie based Instant Blue dye (Expedeon). Gels were submerged in the dye overnight before washing with dH_2O.

2.10.3. Western blot

An SDS-PAGE gel (2.10.2) was sandwiched to a prewashed Amersham Hybond Low Fluorescence 0.2 µm PVDF blotting membrane (GE Healthcare) between filter paper and plastic gauze. The sandwich was submerged in Western transfer buffer (recipe in table 2.15) and transfer carried out through application of 100 mV for 20 minutes at 4 ^{o}C.

Component	Amount
Tris Base	1.82 g
Glycine	8.65 g
Methanol	80 mL
dH_2O	Up to 600 mL
pH adjusted to 8.3-8.4	

Table 2.15: Recipe for Western transfer buffer.

The resultant membrane was washed with TTBS buffer (recipe in table 2.16) and blocked overnight in TTBS containing 0.5 % BSA.

Component	Amount
Tris.HCl	2.42 g
NaCl	29.22 g
Tween 20	0.1 % (v/v)
dH_2O	Up to 1 L

Table 2.16: Recipe for TTBS buffer.

After washing thrice with TTBS, primary antibody was applied in a dilution of 1:10,000 in TTBS with 0.5 % BSA over 1 hour at 4 ^{o}C. The membrane was washed thrice with TTBS before incubation in the presence of 1:5,000 secondary antibody in TTBS with 0.5 % BSA for 1 hour at 4 ^{o}C. The membrane was then washed in TTBS overnight. Both primary and secondary antibodies are described in table 2.1.

To image the western, the membrane was incubated in CDP-Star Chemiluminescent substrate for 5 minutes before imaging with a ChemiDoc-MP (Biorad) using the auto-optimise procedure.

2.10.4. Circular dichroism

Protein buffers were created by titrating 0.1 M solutions of buffer salts to the desired pH as described in table 2.17. Purified protein was transferred to the appropriate pH buffer through two rounds of buffer exchange using Amicon 5000mw concentrator columns. Protein was diluted to 0.1 mg/mL before loading into 1 mm path length quartz cuvettes. Polarised light rotation was measured using a Model 410 Circular Dichroism Spectrophotometer (AVIV), scanning between 190 and 250 nm thrice and calculating the average and standard error. A buffer reading at each pH was read and subtracted from the appropriate trace.

Desired pH	Buffer salts titrated (0.1 M)
2.0 and 3.0	Sodium phosphate monobasic and phosphoric acid
4.0 and 5.0	Sodium acetate and acetic acid
6.0 and 7.0	Sodium phosphate dibasic and sodium phosphate monobasic

Table 2.17: Buffer salts and acids used to create the protein buffers of desired pH.

2.11. Bioinformatics

2.11.1. Using Mega7 and BEAST to generate phylogenetic trees

For both the alignment of peptide sequences and tree building with neighbour joining, maximum likelihood or parsimony algorithms, the program MEGA7 (MegaSoftware) was used.

2.11.1.1. Generating an alignment using ClustalW or Muscle

Sequences were sourced from NCBI and transferred to the “Alignment” tool in MEGA7. Alignments generated through ClustalW and Muscle were performed using the parameters specified in figure 2.6. Positions containing gaps in one or more taxa were removed for all sequences before phylogenetic trees were built. To replicate the alignment that was used in Bialek *et al.* (2008), the sequences were manually aligned used the “Alignment” tool in MEGA7.

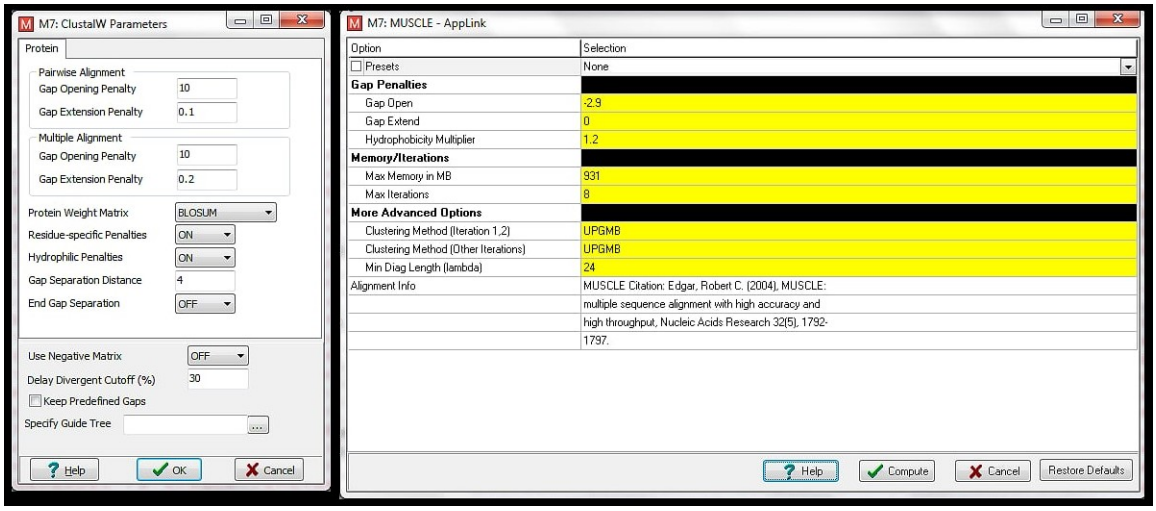


Figure 2.6: Parameters used in MEGA7 for ClustalW (left) or Muscle (right) peptide alignment.

2.11.1.2. Generating phylogenetic trees using Neighbor-Joining, Maximum Likelihood or Parsimony algorithms

Aligned sequences were exported in the “meg.” format and then loaded into MEGA7. The Overall Mean Distance was then calculated by selecting the “Distances” tab followed by “Compute Overall Mean” in the resultant menu. Alignment files were then used to build a phylogenetic tree by selecting the “Phylogeny” tab followed by either “compute phylogenetic tree through Neighbor-Joining", “compute phylogenetic tree through Maximum-likelihood" or “compute phylogenetic tree through Parsimony". The parameters used for Neighbor-Joining and Parsimony are shown in Figure 2.7.

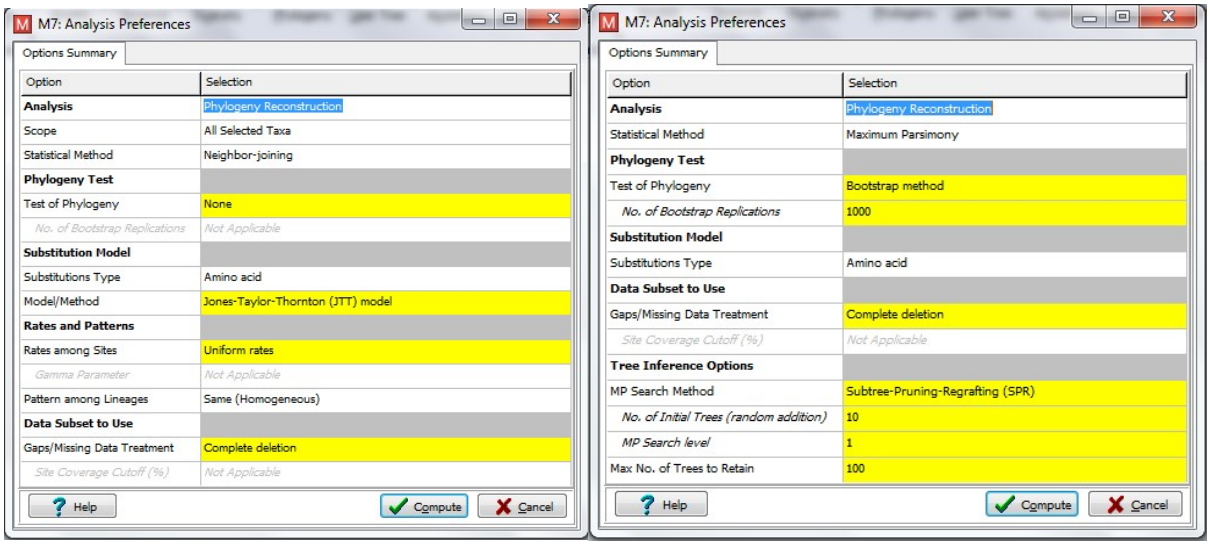


Figure 2.7: Parameters used in the Neighbor-Joining (left) and Parsimony (right) phylogenetic tree building algorithms.

Prior to building a Maximum-likelihood tree, optimal model selection was performed by selecting the "Models" tab and selecting "Find the best DNA/Protein models (ML)". The parameters used for this selection are shown in figure 2.8.

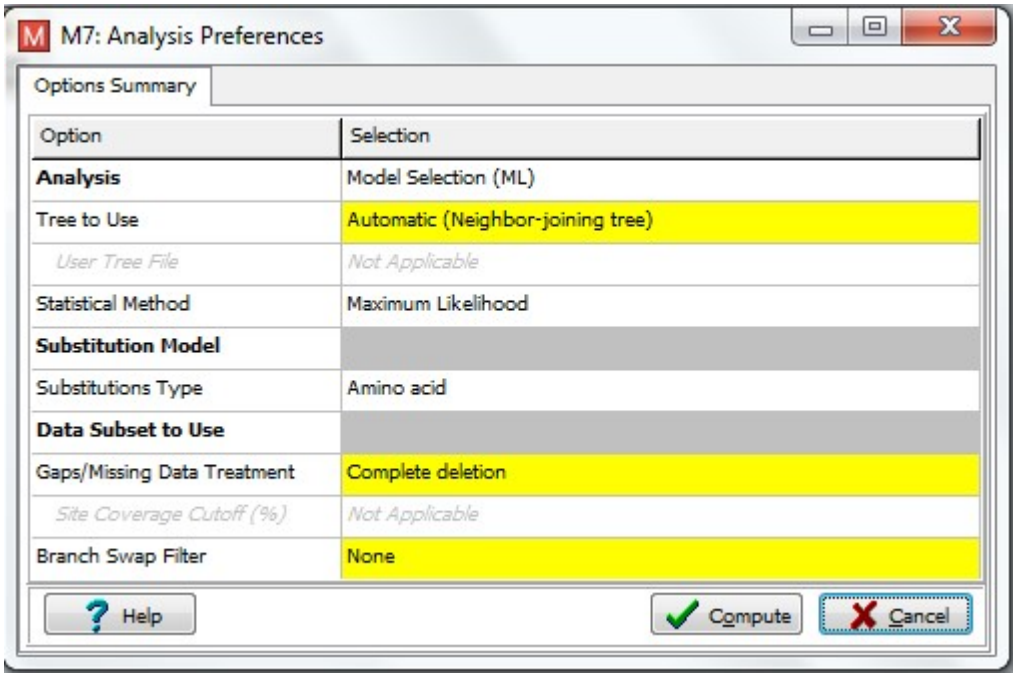


Figure 2.8: Parameters used for finding the optimal model to use in Maximum-Likelihood.

Once the optimal model was found for the alignment, the relevant model was put into the parameters for Maximum-likelihood (default shown in figure 2.9). The "Model/Method" and "Rates among Sites" options were affected by the model selection analysis.

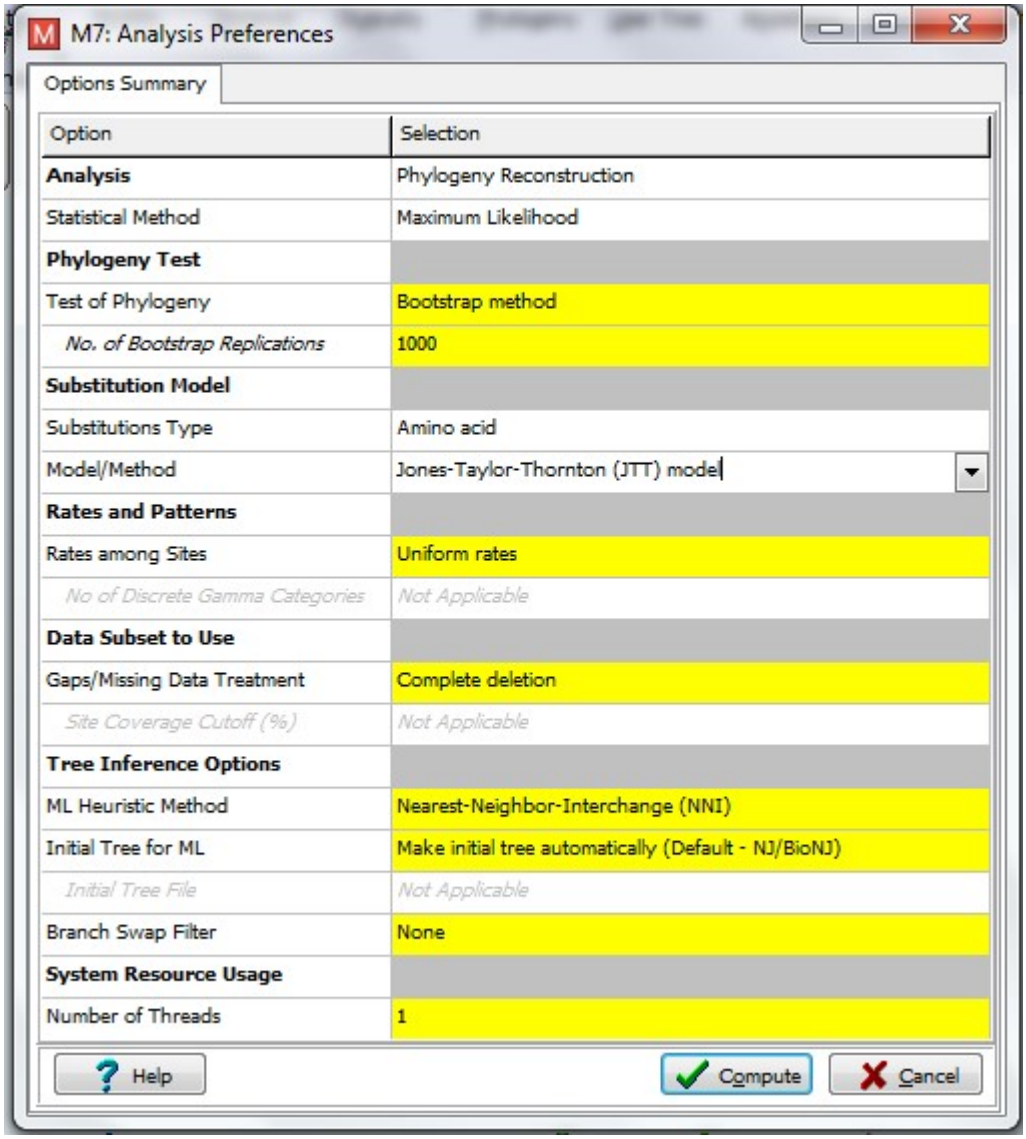


Figure 2.9: Default parameters used in the Maximum-likelihood algorithm for building phylogenetic trees.

2.11.1.3. Generating phylogenetic trees using Bayesian inference

For tree building with Bayesian inference, the program Mega7 (MegaSoftware) was used in conjunction with BEAST. The alignment was exported as a FASTA file before conversion to NEXUS format using FastaConvert. The alignment file was imported into BEAUti (BEAST package) and the parameters set are shown in figure 2.10. Every other parameter was set at default and a BEAST file was generated.

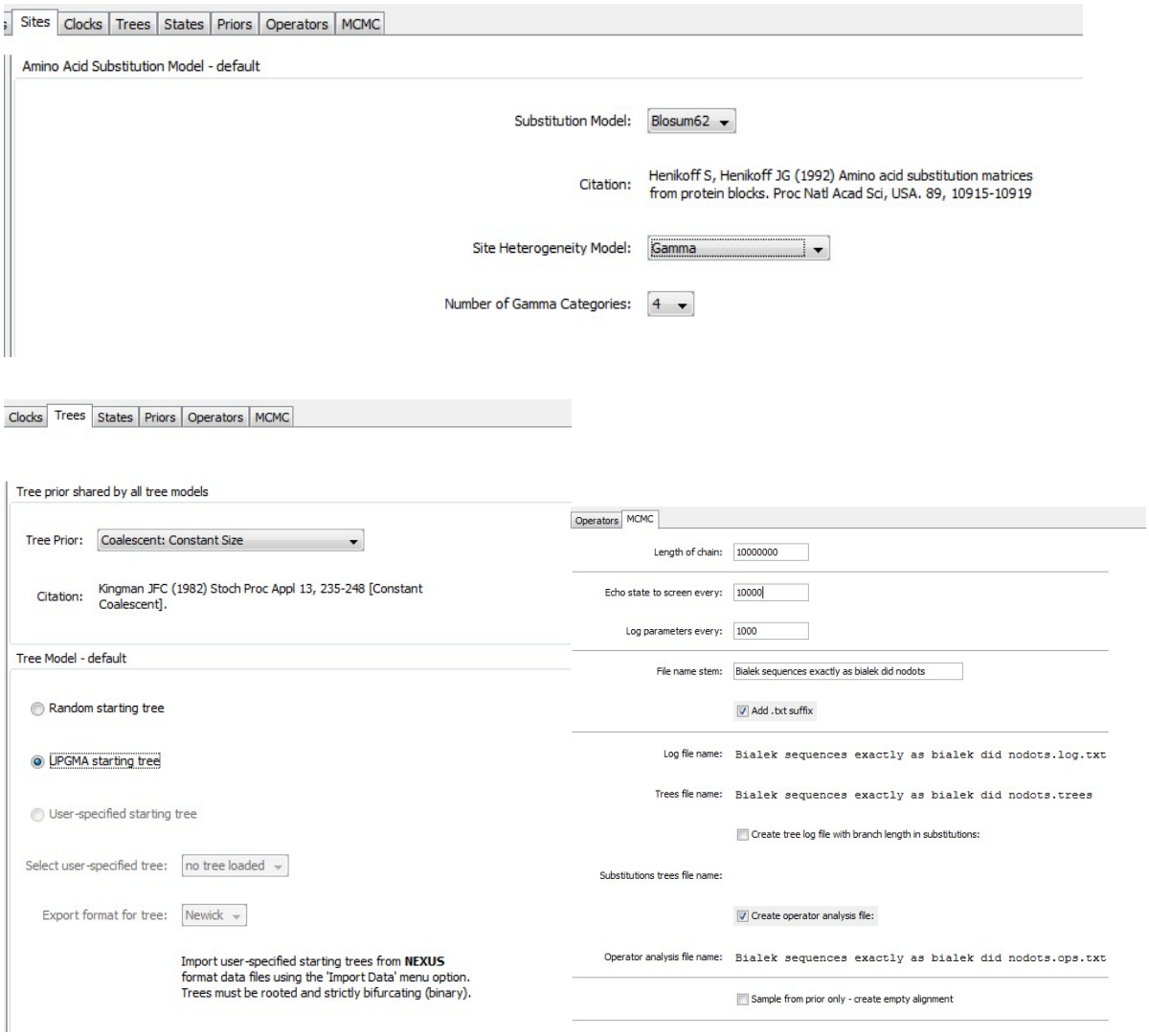


Figure 2.10: Parameters used in BEAUTi to set up the Bayesian Inference of Trees run.

The BEAST file was then opened in BEAST and run with default parameters. The file produced after the completed run was opened in Tracer (BEAST package), and the optimal Burn-In number to exclude the initial low probability trees was determined and used in TreeAnnotator (BEAST package) under the "Specify the burnin as the number of states" option. The final tree formed was visualised and exported to pdf using FigTree (BEAST package).

2.11.2. Using NCBI BLAST

Searches for homologous protein sequences were performed using NCBI BLAST (https://www.ncbi.nlm.nih.gov/). The cytochrome c_6 c_{6A}, c_{6B} and c_{6C} peptide sequences (without signal peptides) used in both BLASTp and tBLASTn searches are listed in table 2.18.

Protein	Organism	Accession
Cytochrome c_6	*Synechocystis* sp. PCC 6803	ALJ67080.1
Cytochrome c_{6A}	*Arabidopsis thaliana*	AED95193.1
Cytochrome c_{6B}	*Prochlorococcus marinus* PCC 1375	AAP99622.1
Cytochrome c_{6C}	*Synechococcus* sp. PCC 7002	ACB00369.1

Table 2.18: Cytochrome c_6 family proteins used in NCBI BLAST searches.

For BLASTp searches the database searched was "Non-redundant protein sequences (nr)" with default parameters. For tBLASTn searches the database searched was "nucleotide collection (nr/nt)" with default parameters. The taxonomic IDs of the clades investigated are listed in table 2.19.

Clade	Taxonomic ID
Embryophyta	3193
Chlorophyta	3041
Glaucophyta	38254
Red Algae	2763
Cryptomonads	3027
Haptophytes	2830
Dinoflagellates	2864
Heterokonts	33634
Rhizaria	543769
Euglenids	3035
Cyanobacteria	1117
Purple phototrophic bacteria	1224
Green sulphur bacteria	1090

Table 2.19: Taxonomic IDs used in NCBI BLAST searches.

3. Phylogenetics and ancestry of the cytochrome c_6 family

3.1. Introduction

3.1.1. Cytochromes c_{6B} and c_{6C}

3.1.1.1. Two homologues of cytochrome c_{6A} were found in cyanobacteria

The search for novel cytochromes has extended beyond eukaryotic photosynthetic organisms to include cyanobacteria. Southern blot screening of the *Synechococcus sp.* PCC 7002 genome revealed a homologue of cytochrome c_6 (PETJ) subsequently designated PETJ2 (Nomura, 2001). In the same study the *petJ2* gene and peripheral regions were sequenced, and in a later study PETJ2 was cloned into a pUC plasmid (pUCpetJ2), expressed and revealed to have a significantly lower redox midpoint potential than PETJ from *Synechococcus sp.* PCC 7002 (+155 mV and +319 mV respectively) and a more positive surface charge (Bialek *et al.* 2008, 2014), indicating that PETJ2 is unable to accept electrons from cytochrome b_6f. Therefore, like cytochrome c_{6A}, PETJ2 is highly unlikely to be a functional substitute for cytochrome c_6. This alteration in redox midpoint potential was also found to be due to the haem environment of the cytochrome, with the presence of a hydrophobic residue at around amino acid 50 contributing strongly to the lower value (Bialek *et al.* 2014). Unlike cytochrome c_{6A}, however (which has not been found in cyanobacteria), PETJ2 does not have a LIP sequence containing a disulphide bridge.

Phylogenetic analysis of the PETJ2 amino acid sequence in a range of cyanobacterial species suggested the existence of two distinct groups of cytochrome c_6 homologues in cyanobacteria, tentatively named cytochromes c_{6B} and c_{6C} (Bialek *et al.*, 2008). A multiple alignment of the peptide sequences of various c-type cytochromes (c_6, c_{6A}, c_{6B}, c_{6C}, c_M, c_{550} and c_{555}) across both eukaryotic and prokaryotic photosynthetic organisms was performed using ClustalX v 1.83.1, and the subsequent alignment used by Bialek *et al.* (2008) to build an unrooted phylogenetic tree using Neighbor-joining in PHYLIP (Felsenstein, 1989; figure 3.1).

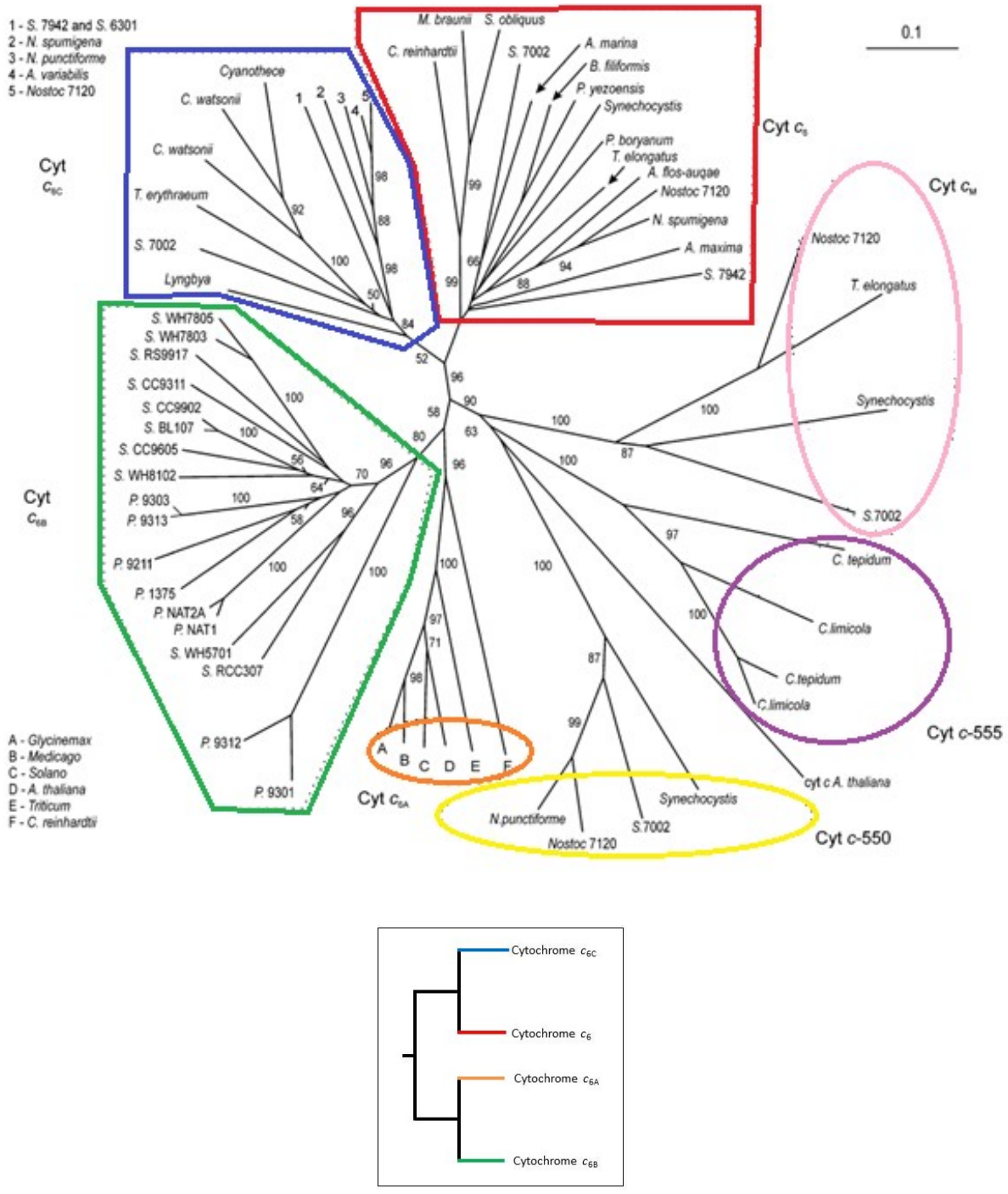


Figure 3.1: Radial tree of c-type cytochromes built using NJ from a ClustalX peptide alignment. Coloured boxes surround clusters designated with the relevant c-type cytochrome, with blue, red, pink, purple, yellow, orange and green representing cytochromes c_{6C}, c_6, c_M, c-555, c-550, c_{6A} and c_{6B} respectively. The bar at the top right denotes an approximate evolutionary distance (substitutions per residue). Bootstrap values above 50 are shown by the relevant branch point. In the cytochrome c_{6B} cluster S stands for Synechococcus and P stands for Prochlorococcus. The rooted tree in the box at the bottom represents the proposed simplified ancestry of the c_6-type cytochromes. Adapted from Bialek et al., 2008.

The cytochromes c_{6B} and c_{6C} were identified and segregated based on the deep-branching clusters seen in the phylogenetic tree. The cytochrome c_{6B} cluster appeared to share a more recent ancestry with cytochrome c_{6A} than with the other proteins, whereas cytochrome c_{6C} appeared to share a more recent ancestry with cytochrome c_6 (figure 3.1, black box). The authors also noted that cytochrome c_{6B} was predominantly found in the marine species of *Prochlorococcus* and *Synechococcus*, which are unable to fix nitrogen, whereas cytochrome c_{6C} was predominantly found in nitrogen-fixing cyanobacteria.

Crystal structures have been determined and functional analysis performed on cytochromes c_{6B} and c_{6C} (figure 3.2). X-ray crystallography performed on cytochromes c_{6A}, c_{6B} and c_{6C} shows a very similar overall structure, with 4 alpha helices linked with loops interacting with a haem prosthetic group. The positioning of the alpha helices in cytochromes c_{6B} and c_{6C} appears very similar, especially when compared to those of cytochrome c_{6A}, and the lack of a LIP from cytochromes c_{6B} and c_{6C} is notable. The amino acid aligned to position 52 in *A. thaliana* cytochrome c_{6A} in figure 3.2 is also notable, as the nature of this amino acid has been found to contribute greatly to the lower redox midpoint potential of cytochromes $c_{6A/B/C}$ (hydrophobic) compared to cytochrome c_6 (glutamine; Worrel *et al.*, 2007; Bialek *et al.*, 2008).

Characterisation of cytochrome c_{6C} knockout mutant lines of *Synechococcus* PCC 7942 showed diminished growth rate relative to wild type lines under intense light conditions, suggesting a role in high light stress response (Vasudevan, PhD thesis, University of Cambridge). These data combined with transcriptomic analysis of the lines under different light conditions, whereby alternative electron transfer pathways through cytochrome c_6 and c_M were upregulated relative to wild type at high light intensities, suggested a role transferring electrons from the plastoquinone pool to plastocyanin or PSI, thus bypassing the bottleneck of electron transfer through cytochrome b_6f and decreasing ROS production (Vasudevan, unpublished). This was supported by chlorofluorescence measurements performed on the c_{6C} knockout lines, which demonstrated a lower level of non-photochemical quenching (NPQ) relative to the wild type line that suggested a decrease in cyclic electron flow which relies on connecting PQ to PSI (Vasudevan, unpublished). No similar functional studies have yet been performed on cytochrome c_{6B}.

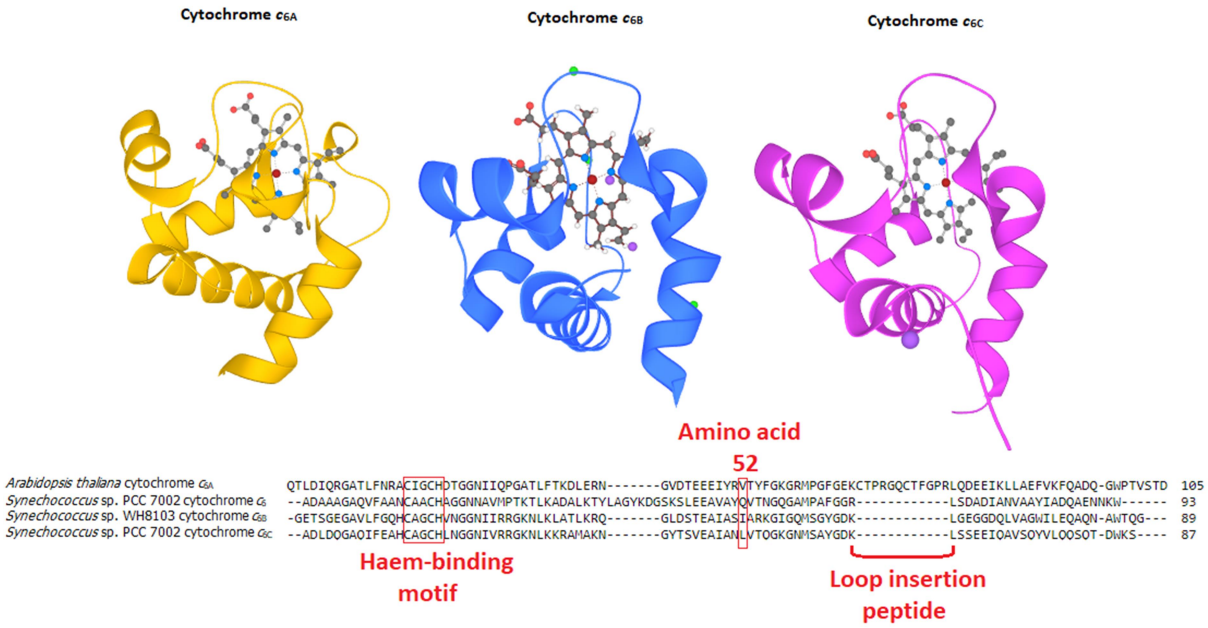


Figure 3.2: X-ray crystal structures and peptide sequence alignments of cytochromes c_{6A} (yellow), c_{6B} (blue) and c_{6C} (magenta). Secondary structure for each protein is shown in ribbon form, and the haem prosthetic groups shown in ball and stick (carbon-black, oxygen- pink, nitrogen- blue and iron-deep red). Peptide alignments were performed using Clustal Omega (Chojnacki et al., 2017). The haem-binding motif (CXXCH), the LIP, and amino acid 52 have been highlighted in red. Figure uses crystallography and sequence data from Marcaida et al., 2006; Zatwarnicki et al., 2014; Bialek et al., 2014 (Q93VA3, Q7U624 and Q8KX15 on Uniprot respectively).

3.1.1.2. The distinction between cytochromes c_{6B} and c_{6C} warrants further investigation

An essential question from this discovery is if the difference observed between cytochromes c_{6B} and c_{6C} is genuine, and not an artefact created from the strategy for phylogenetic inference used. For example, the sequences chosen for cytochrome c_{6B} had been sourced from subspecies of only two genera of cyanobacteria (*Prochlorococcus* and *Synechococcus*), yet the sequences chosen for cytochrome c_{6C} had been sourced from a wider range of cyanobacteria. This could result in the two groups of sequences being separated in a similar arrangement as the species would be based on sequences of the same conserved gene. This would imply that cytochrome c_{6B} and c_{6C} are actually functionally the same, and that the clusters observed are due to the evolutionary relationship of the species themselves. Subsequently, if the phylogenetic difference were confirmed between the two clusters, then the question would remain as to whether the two sets of proteins are functionally

different, or whether they are orthologues that perform the same role in two distinct clades of cyanobacteria.

The clusters observed in the phylogenetic tree do suggest that c_{6C} shows most recent common ancestry with c_6 and c_{6B} shows most recent common ancestry with c_{6A} (figure 3.1). However, according to multiple alignments in BLAST performed by the researchers, the PETJ2 mature sequence from *Synechococcus sp.* PCC 7002 shows similar identities and similarities to both c_6 and c_{6B} (43.5/60.1% and 44.6/64.4% respectively, Bialek *et al.*, 2008). These values might suggest that cytochrome c_{6C} is marginally more closely related to c_{6B} than c_6, and therefore that the two novel cytochromes may have a similar function, although the values are similar, and percent similarities do not necessarily reflect evolutionary relationships accurately.

The bootstrap values and scope of the phylogenetic tree also throw doubt onto the distinction between cytochromes c_{6B} and c_{6C}. Bootstrap values below 50 have been excluded from the tree, therefore hiding weaker connecting branches, and the edge shared by cytochromes c_{6B} and c_{6A} has a low bootstrap value of 58, which is arguably too low to be considered reliable. The inclusion of the long edges to cytochromes c_M, c_{550} and c_{555} may also be affecting the results of the tree-building and are unnecessary as the cytochromes discussed are not members of the c_6 family and perform a different role from them (Kusaj and Yamanaka, 1973; Shen and Inoue, 1993; Vasudevan, personal communication). Therefore, the tree may be improved in robustness through the removal of these superfluous sequences from the study. The tree may also benefit from including more sequences from a wider range of cyanobacteria and other photosynthetic organisms containing c_6-type cytochromes (section 3.1.2). This wider scope of organisms could help determine not only whether the cytochrome c_{6B} and c_{6C} segregation is robust, but also help build a stronger model of the ancestry and evolution of all the c_6-type cytochromes.

3.1.2. A model has been proposed for the ancestry and evolution of the cytochrome c_6 family

The current hypothesis for how the cytochrome c_6 family evolved in photosynthetic organisms has been outlined by Howe *et al.* (2015; Figure 3.3). The presence of plastocyanin, cytochrome c_6 and members of the cytochromes $c_{6A}/c_{6B}/c_{6C}$ family in current cyanobacteria and some eukaryotic lineages implies that ancestral cyanobacteria, including those destined to undergo endosymbiosis and become plastids, also contained these genes (or representatives of the gene families). After endosymbiosis, the green plant lineage (including higher plants and unicellular green algae) retained a form of cytochrome c_{6A}, c_{6B} or c_{6C} (hereby collectively termed $c_{6A/B/C}$), although higher plants lost

cytochrome c_6 and *Euglena* lost cytochrome $c_{6A/B/C}$ after secondary endosymbiosis. The glaucophytes, red algae and haptophytes appear to have lost plastocyanin, and glaucophytes and rhodophytes subsequently lost cytochrome $c_{6A/B/C}$, although the retention of $c_{6A/B/C}$ in haptophytes, which have a secondary plastid of red algal origin, implies this occurred post-secondary endosymbiosis. This model therefore suggests that gene duplication led to the creation of cytochrome c_6 and $c_{6A/B/C}$ in the cyanobacterial ancestor to the chloroplast ancestral chloroplast, where the modifications leading to differing redox midpoint potentials and surface charge would have occurred.

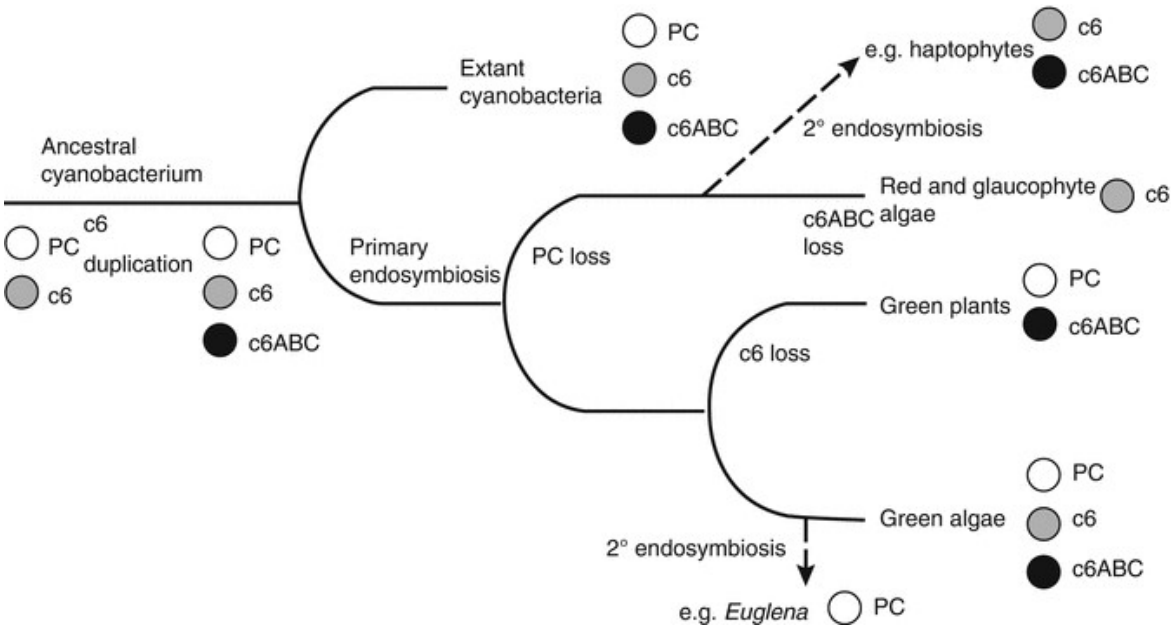


Figure 3.3: Current model of cytochrome c_6 family evolution in photosynthetic organisms. PC stands for plastocyanin. Taken from Howe et al., 2015.

3.1.2.1. Further refinement can be applied to the model

There are some assumptions in the model that can be addressed using newly discovered information. Firstly, the sequencing and annotation of more genomes of photosynthetic organisms means that a wider and more detailed search for cytochrome $c_{6A/B/C}$ family genes can be performed and added to the model. It is also noticeable that the cytochrome $c_{6A/B/C}$ family have been all represented together, despite the observation that cytochrome c_{6A} has a LIP where cytochromes c_{6B} and c_{6C} do not. Cytochrome c_{6A} was also identified only in eukaryotes, whereas the cytochromes c_{6B} and c_{6C} have been identified only in prokaryotes. Therefore, by understanding the evolution of the members of the cytochrome $c_{6A/B/C}$ subfamily the model of how cytochrome c_{6A} evolved (e.g. when the LIP was inserted) will become clearer.

3.1.3. Certain genes have been found to be closely linked to cytochrome c_{6A} in *Arabidopsis thaliana*

In *A. thaliana* two genes located near the coding region for cytochrome c_{6A} (*petJ*) have been noted as particularly interesting. One is a potential trypsin-family protease (AT5G45030) which has overlapping 3' UTRs with *petJ* (AT5G45040; figure 3.4). This overlapping of non-coding regions may suggest that expression of the two genes may be coordinated, which could in turn suggest coordinated function.

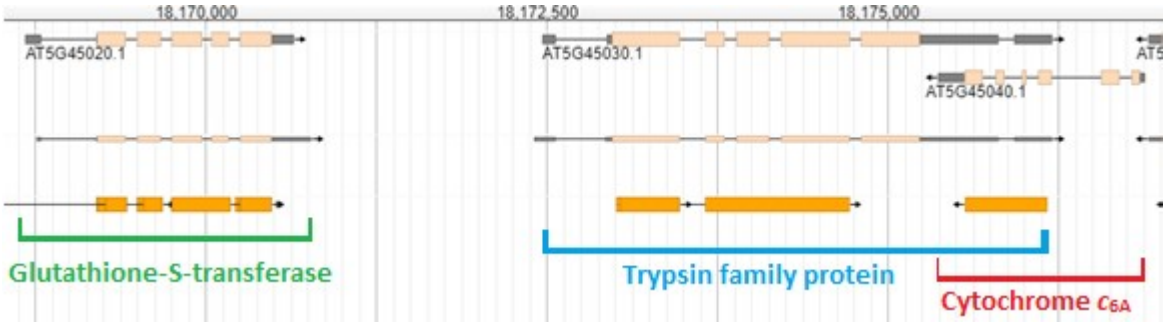


Figure 3.4: Coding region of cytochrome c_{6A} in A. thaliana genome on chromosome 5. Cytochrome c_{6A} (AT5G45040.1, red), Trypsin family protein (AT5G45030.1, blue) and glutathione-S-transferase protein (AT5G45020.1. green) coding regions are highlighted. The numbers at the top refer to nucleotide position on chromosome 5, and the beige rectangles on the gene represent exons, with the line in between representing introns. The grey rectangles flanking the genes are UTRs.

The other gene found ~4.5 kb downstream of *petJ* encodes a glutathione-S-transferase family protein (AT5G45020; figure 3.4). This close linkage could be of interest as glutathione-S-transferases catalyse redox reactions involving thiol molecules, suggesting the hypothetical protein could be involved in LIP redox chemistry in cytochrome c_{6A}.

If this close linkage between cytochrome c_{6A} and these two genes is conserved across higher plants then it may suggest a coordinated function. This would suggest a need for further characterisation of the two genes of interest, as at the time of writing the precise function of either is unknown.

3.1.4. Aims of the chapter

The aim of this chapter is to use bioinformatic techniques to achieve the following goals:

- To recreate the phylogenetic tree used by Bialek *et al.* (2008) to suggest two separate subfamilies of cytochrome c_6 (c_{6B} and c_{6C}) in order to test its robustness.
- To create a new phylogenetic tree using cytochrome c_6 family sequences from a wider range of photosynthetic organisms to see if a similar clustering between c_{6B} and c_{6C} occurs. This tree will also help to refine the evolutionary model of c_6 family proteins.
- To refine further the evolutionary model of c_6 family proteins by scanning proteomic data of a wider range of photosynthetic organisms to find more homologues of cytochromes $c_{6A/B/C}$.
- To assess how conserved the close linkage observed between cytochrome c_{6A} and Trypsin family protein and glutathione-S-transferase is across the green eukaryotic lineage.

3.2. Results

3.2.1. Is there a phylogenetic difference between cytochromes c_{6B} and c_{6C}?

3.2.1.1. Recreating the phylogenetic tree segregating cytochromes c_{6B} and c_{6C} using differing alignment and tree-building algorithms

Protein sequences for all the organisms analysed in the phylogenetic tree identifying cytochromes c_{6B} and c_{6C} were sourced from the supplementary information of Bialek *et al.* (2008). The sequences had already been processed to remove non-coding regions and signal peptides. The sequences were loaded into MEGA7 (Kumar *et al.*, 2016) and aligned using either the ClustalW algorithm (Thompson *et al.*, 2003) or the Muscle algorithm (Edgar 2004). The ClustalX alignment reported by Bialek *et al*. (2008, supplementary information) was also used, and will be referred to as the Bialek-ClustalX alignment. The Overall Mean Distance (OMD) was then calculated for all three alignments (figure 3.5). The OMD gives an idea of the overall spread of the alignment through calculating the mean pairwise distance and standard error for each alignment, which can then give an indication as to which phylogenetic tree building algorithm is most appropriate, with a value greater than 1 showing that the data are too spread out for Neighbor-Joining (NJ) to be appropriate (Hall, 2017).

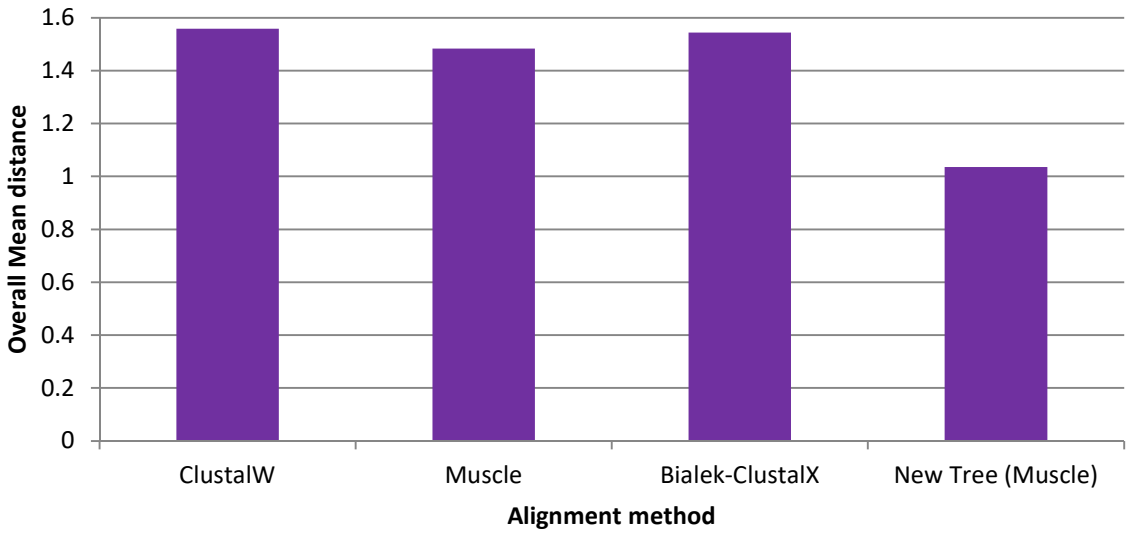
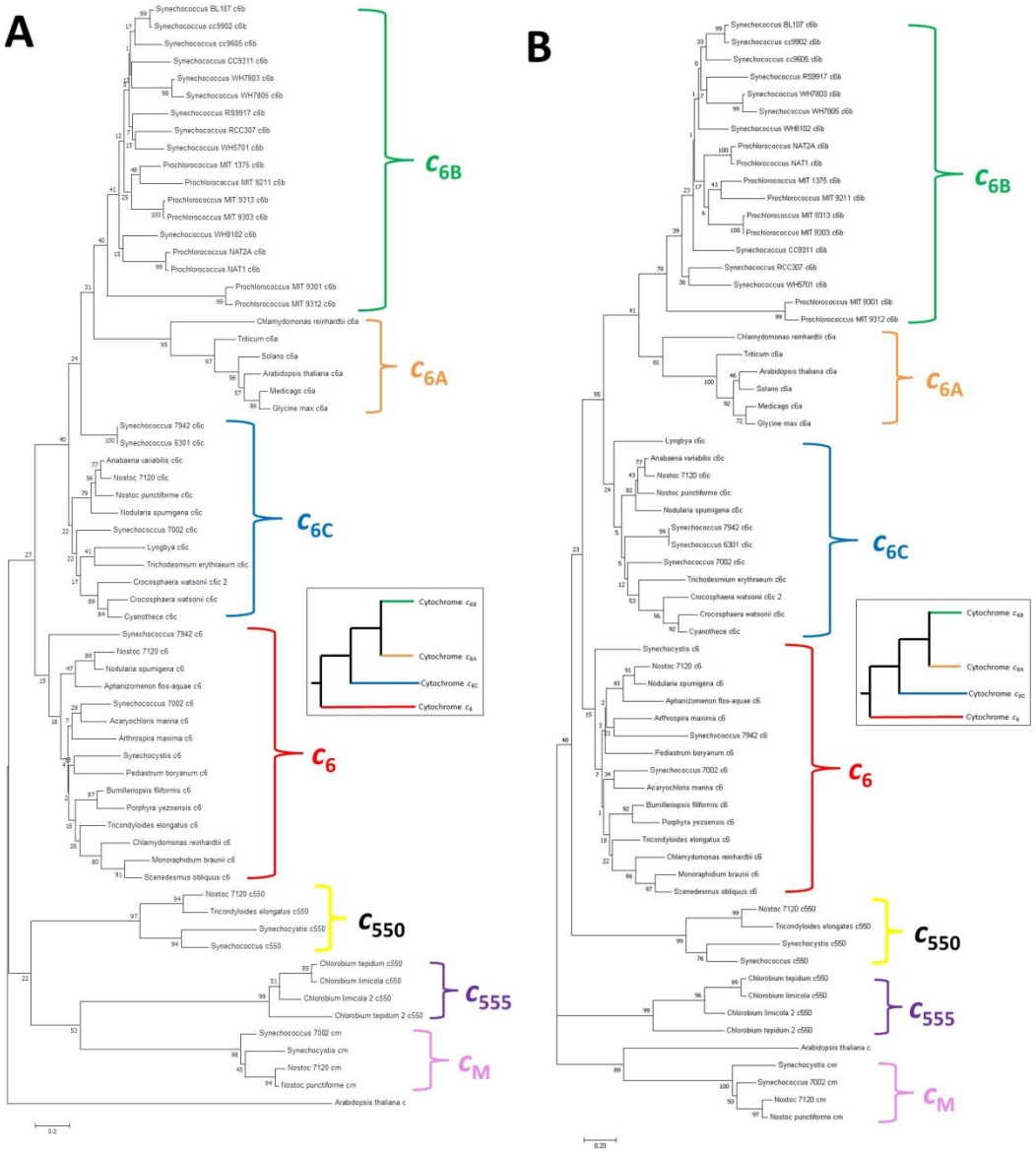


Figure 3.5: Overall Mean Distances calculated for each alignment used in this study. ClustalW and Muscle refer to the alignments performed on the original sequences from Bialek et al. (2008), with 'Bialek-ClustalX' refering to the alignment used in Bialek et al. (2018). New Tree (Muscle) refers to the OMD calculated for the alignment created for the new selection of sequences built in 3.2.1.2.

The ClustalW, Muscle and Bialek-ClustalX alignments were then used to build phylogenetic trees using NJ (figure 3.6). This tree-building method was used because it was the one used in the original paper. The distribution of bootstrap values observed in the trees was noticeably wider than that of the original tree (see figure 3.1), and the bootstrap values had a lower average value (figure 3.7).


A
Synechococcus BL107 c6b
Synechococcus cc9902 c6b
Synechococcus cc9605 c6b
Synechococcus CC9311 c6b
Synechococcus WH7803 c6b
Synechococcus WH7805 c6b
Synechococcus RS9917 c6b
Synechococcus RCC307 c6b
Synechococcus WH5701 c6b
Prochlorococcus MIT 1375 c6b
Prochlorococcus MIT 9211 c6b
Prochlorococcus MIT 9313 c6b
Prochlorococcus MIT 9303 c6b
Synechococcus WH8102 c6b
Prochlorococcus NAT2A c6b
Prochlorococcus NAT1 c6b
Prochlorococcus MIT 9301 c6b
Prochlorococcus MIT 9312 c6b
Chlamydomonas reinhardtii c6a
Triticum c6a
Solano c6a
Arabidopsis thaliana c6a
Medicago c6a
Glycine max c6a
Synechococcus 7942 c6c
Synechococcus 6301 c6c
Anabaena variabilis c6c
Nostoc 7120 c6c
Nostoc punctiforme c6c
Nodularia spumigena c6c
Synechococcus 7002 c6c
Lyngbya c6c
Trichodesmium erythraeum c6c
Crocosphaera watsonii c6c 2
Crocosphaera watsonii c6c
Cyanothece c6c
Synechococcus 7942 c6
Nostoc 7120 c6
Nodularia spumigena c6
Aphanizomenon flos-aquae c6
Synechococcus 7002 c6
Acaryochloris marina c6
Arthrospira maxima c6
Synechocystis c6
Pediastrum boryanum c6
Bumilleriopsis filiformis c6
Porphyra yezoensis c6
Tricondyloides elongatus c6
Chlamydomonas reinhardtii c6
Monoraphidium braunii c6
Scenedesmus obliquus c6
Nostoc 7120 c550
Tricondyloides elongatus c550
Synechocystis c550
Synechococcus c550
Chlorobium tepidum c550
Chlorobium limicola c550
Chlorobium limicola 2 c550
Chlorobium tepidum 2 c550
Synechococcus 7002 cm
Synechocystis cm
Nostoc 7120 cm
Nostoc punctiforme cm
Arabidopsis thaliana c
0.2
Cytochrome c6B
Cytochrome c6A
Cytochrome cCC
Cytochrome c6
c6B
c6A
c6C
c6
c550
c555
cM
B
Synechococcus BL107 c6b
Synechococcus cc9902 c6b
Synechococcus cc9605 c6b
Synechococcus RS9917 c6b
Synechococcus WH7803 c6b
Synechococcus WH7805 c6b
Synechococcus WH8102 c6b
Prochlorococcus NAT2A c6b
Prochlorococcus NAT1 c6b
Prochlorococcus MIT 1375 c6b
Prochlorococcus MIT 9211 c6b
Prochlorococcus MIT 9313 c6b
Prochlorococcus MIT 9303 c6b
Synechococcus CC9311 c6b
Synechococcus RCC307 c6b
Synechococcus WH5701 c6b
Prochlorococcus MIT 9301 c6b
Prochlorococcus MIT 9312 c6b
Chlamydomonas reinhardtii c6a
Triticum c6a
Arabidopsis thaliana c6a
Solano c6a
Medicago c6a
Glycine max c6a
Lyngbya c6c
Anabaena variabilis c6c
Nostoc 7120 c6c
Nostoc punctiforme c6c
Nodularia spumigena c6c
Synechococcus 7942 c6c
Synechococcus 6301 c6c
Synechococcus 7002 c6c
Trichodesmium erythraeum c6c
Crocosphaera watsonii c6c 2
Crocosphaera watsonii c6c
Cyanothece c6c
Synechocystis c6
Nostoc 7120 c6
Nodularia spumigena c6
Aphanizomenon flos-aquae c6
Arthrospira maxima c6
Synechococcus 7942 c6
Pediastrum boryanum c6
Synechococcus 7002 c6
Acaryochloris marina c6
Bumilleriopsis filiformis c6
Porphyra yezoensis c6
Tricondyloides elongatus c6
Chlamydomonas reinhardtii c6
Monoraphidium braunii c6
Scenedesmus obliquus c6
Nostoc 7120 c550
Tricondyloides elongatus c550
Synechocystis c550
Synechococcus c550
Chlorobium tepidum c550
Chlorobium limicola c550
Chlorobium limicola 2 c550
Chlorobium tepidum 2 c550
Arabidopsis thaliana c
Synechocystis cm
Synechococcus 7002 cm
Nostoc 7120 cm
Nostoc punctiforme cm
0.20
Cytochrome c6B
Cytochrome c6A
Cytochrome cCC
Cytochrome c6
c6B
c6A
c6C
c6
c550
c555
cM

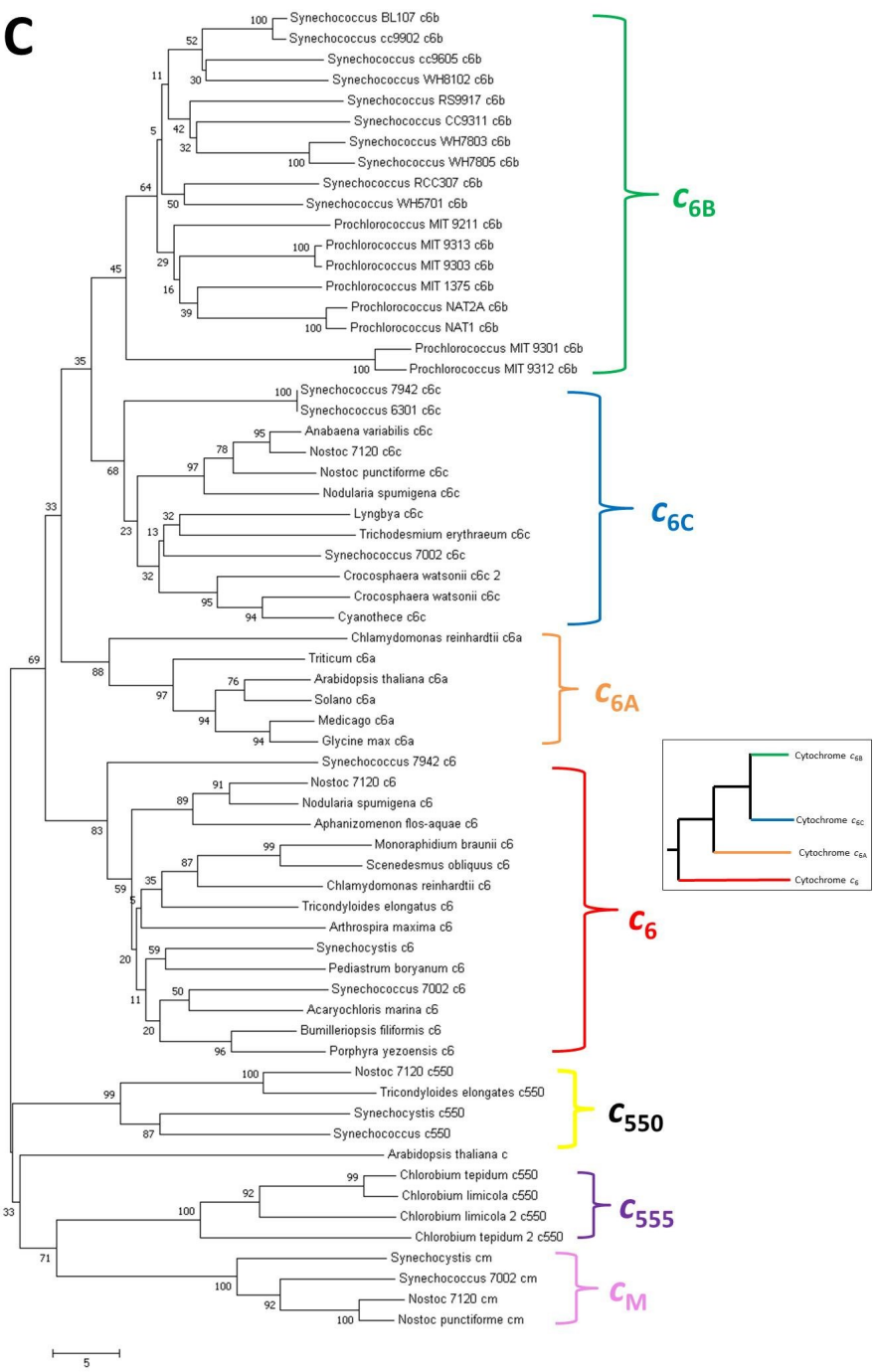


Figure 3.6: Phylogenetic trees built using Neighbor-Joining. Tree A was built from the ClustalW alignment, whereas Tree B was built from the Muscle alignment and Tree C from the Bialek-ClustalX alignment. The original clusters from the Bialek et al. (2008) paper have been marked in the same colour as seen in figure 3.1. Bootstrap values are given at branchpoints. In the boxes are simplified forms of each tree demonstrating the phylogeny of only the c_6-type cytochromes, with the root placed according to where the cytochromes c_M, c_{550} and c_{555} join the tree.

The individual clusters in each tree created were similar in composition to those observed in the original Bialek *et al.* (2008) ClustalX tree. However, in all three trees cytochromes c_{6A}, c_{6B} and c_{6C} clustered together to the exclusion of c_6, based on a root using cytochromes c_M, c_{550} and $c_{555.}$ In the ClustalW tree cytochromes c_{6A} and c_{6B} formed a monophyletic group within the cytochrome c_{6C} group, whereas in the Muscle tree cytochromes c_{6A} and c_{6B} formed a monophyletic sister group to cytochrome c_{6C}, and in the Bialek alignment tree cytochromes c_{6B} and c_{6C} clustered together as a sister group to cytochrome c_{6A}. These three trees suggest that cytochromes c_{6A} c_{6B} and c_{6C} share common ancestry to the exclusion of cytochrome c_6. This contrasts with the conclusion of Bialek *et al.* (2008) who placed the root between a lineage containing cytochromes c_{6A} and c_{6B} and a lineage containing cytochromes c_6 and $c_{6C.}$ It is not clear why the NJ tree recovered here differed from the one reported by Bialek *et al.* (2008) with the same alignment. This may reflect the low bootstrap values seen, even compared with those from the original paper (figure 3.7), and in any case it is not clear why the bootstrap values should be lower. The tree recovered using the Muscle alignment could be transformed into that recovered by Bialek *et al.* (2008) if the root were placed differently.

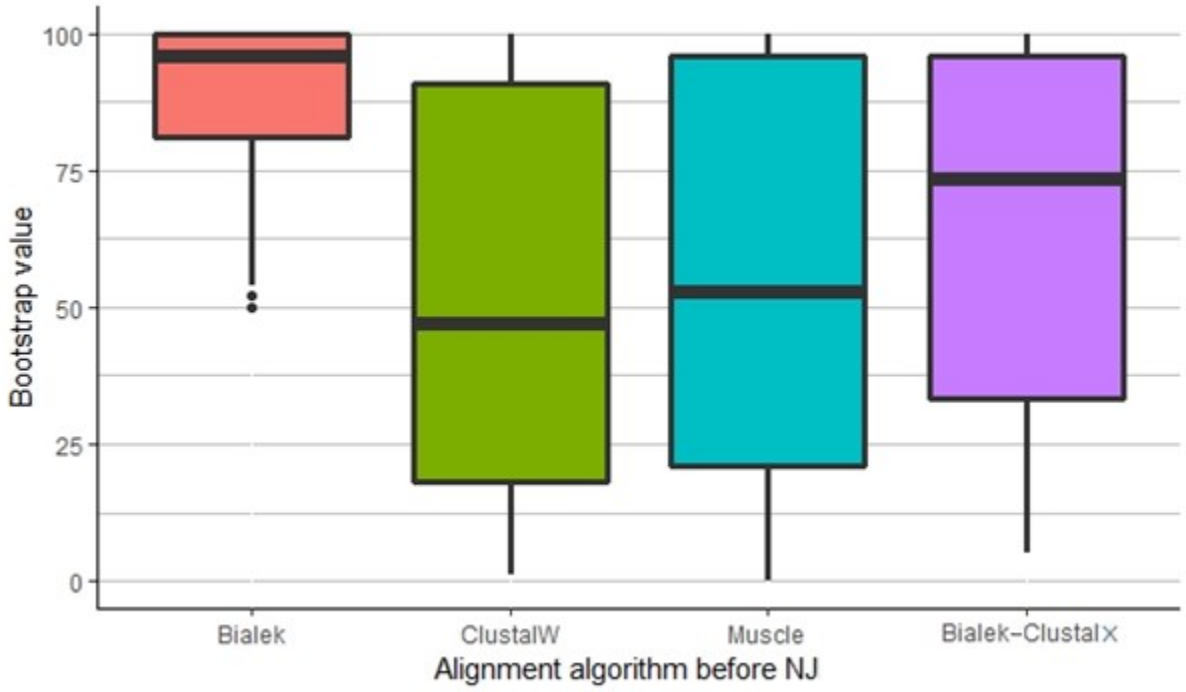
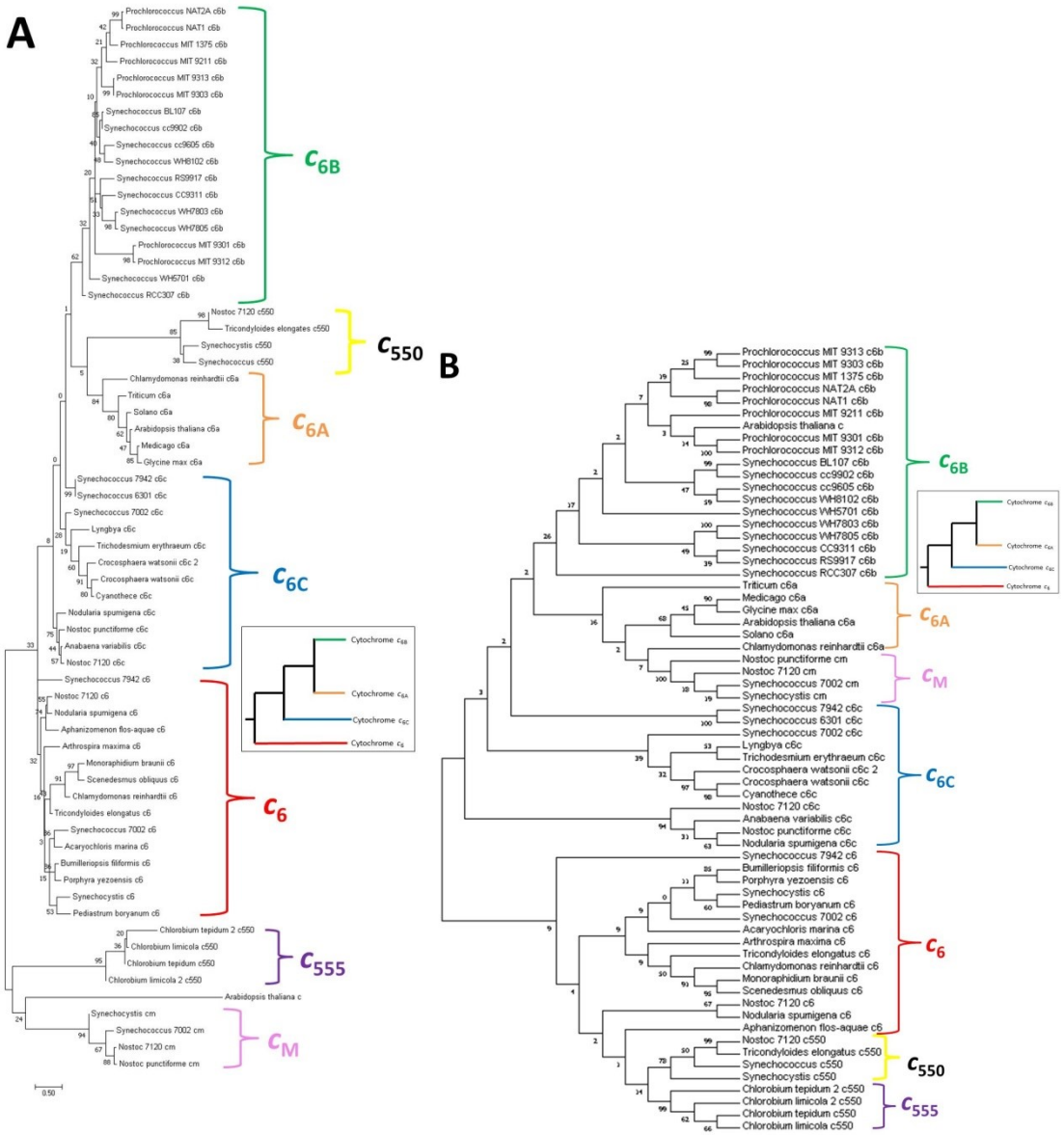


Figure 3.7: Distribution of the bootstrap values recorded from the Bialek et al. (2008), ClustalW, Muscle and Bialek-ClustalX Neighbor-Joining trees.

As the OMD observed for each alignment was found to be over 1, alternative algorithms to NJ were tried to see if similar topologies were recovered. The program MEGA7 was used with the Bialek-ClustalX alignment to create additional phylogenetic trees through Bayesian inference (BI), Maximum Parsimony and Maximum likelihood (ML; figure 3.8). The Maximum likelihood tree used

the Gamma distributed WAG model. The ML and Parsimony trees both had a lower average bootstrap value (figure 3.9), and were difficult to interpret, not least because the outgroup sequences were separated. For example, cytochrome c_{550} clustered with cytochrome c_{6A} in the ML tree (figure 3.8A) as did cytochrome c_M for the Parsimony tree (figure 3.8C). In addition, some of the sequences of cytochrome c_{6C} for example shared a more recent common ancestor with other c_{6A} and c_{6B} sequences than with the majority of c_{6C} sequences (such as those from *Synechococcus* PCC7942 and 6301 cytochrome c_{6C} in the Parsimony tree). Nevertheless, cytochromes $c_{6A,}$ c_{6B} and c_{6C} again appeared to cluster together to the exclusion of c_6, as was seen with the NJ trees built using the alignment sourced from Bialek *et al.* (2008).


A
Prochlorococcus NAT2A c6b
Prochlorococcus NAT1 c6b
Prochlorococcus MIT 1375 c6b
Prochlorococcus MIT 9211 c6b
Prochlorococcus MIT 9313 c6b
Prochlorococcus MIT 9303 c6b
Synechococcus BL107 c6b
Synechococcus cc9902 c6b
Synechococcus cc9605 c6b
Synechococcus WH8102 c6b
Synechococcus RS9917 c6b
Synechococcus CC9311 c6b
Synechococcus WH7803 c6b
Synechococcus WH7805 c6b
Prochlorococcus MIT 9301 c6b
Prochlorococcus MIT 9312 c6b
Synechococcus WH5701 c6b
Synechococcus RCC307 c6b
c_{6B}
Nostoc 7120 c550
Tricondyloides elongates c550
Synechocystis c550
Synechococcus c550
c_{550}
Chlamydomonas reinhardtii c6a
Triticum c6a
Solano c6a
Arabidopsis thaliana c6a
Medicago c6a
Glycine max c6a
c_{6A}
Synechococcus 7942 c6c
Synechococcus 6301 c6c
Synechococcus 7002 c6c
Lyngbya c6c
Trichodesmium erythraeum c6c
Crocosphaera watsonii c6c 2
Crocosphaera watsonii c6c
Cyanothece c6c
Nodularia spumigena c6c
Nostoc punctiforme c6c
Anabaena variabilis c6c
Nostoc 7120 c6c
c_{6C}
Cytochrome c_{6B}
Cytochrome c_{6A}
Cytochrome c_{6C}
Cytochrome c_6
Synechococcus 7942 c6
Nostoc 7120 c6
Nodularia spumigena c6
Aphanizomenon flos-aquae c6
Arthrospira maxima c6
Monoraphidium braunii c6
Scenedesmus obliquus c6
Chlamydomonas reinhardtii c6
Tricondyloides elongatus c6
Synechococcus 7002 c6
Acaryochloris marina c6
Bumilleriopsis filiformis c6
Porphyra yezoensis c6
Synechocystis c6
Pediastrum boryanum c6
c_6
Chlorobium tepidum 2 c550
Chlorobium limicola c550
Chlorobium tepidum c550
Chlorobium limicola 2 c550
c_{555}
Arabidopsis thaliana c
Synechocystis cm
Synechococcus 7002 cm
Nostoc 7120 cm
Nostoc punctiforme cm
c_M
0.50
B
Prochlorococcus MIT 9313 c6b
Prochlorococcus MIT 9303 c6b
Prochlorococcus MIT 1375 c6b
Prochlorococcus NAT2A c6b
Prochlorococcus NAT1 c6b
Prochlorococcus MIT 9211 c6b
Arabidopsis thaliana c
Prochlorococcus MIT 9301 c6b
Prochlorococcus MIT 9312 c6b
Synechococcus BL107 c6b
Synechococcus cc9902 c6b
Synechococcus cc9605 c6b
Synechococcus WH8102 c6b
Synechococcus WH5701 c6b
Synechococcus WH7803 c6b
Synechococcus WH7805 c6b
Synechococcus CC9311 c6b
Synechococcus RS9917 c6b
Synechococcus RCC307 c6b
c_{6B}
Cytochrome c_{6B}
Cytochrome c_{6A}
Cytochrome c_{6C}
Cytochrome c_6
Triticum c6a
Medicago c6a
Glycine max c6a
Arabidopsis thaliana c6a
Solano c6a
Chlamydomonas reinhardtii c6a
c_{6A}
Nostoc punctiforme cm
Nostoc 7120 cm
Synechococcus 7002 cm
Synechocystis cm
c_M
Synechococcus 7942 c6c
Synechococcus 6301 c6c
Synechococcus 7002 c6c
Lyngbya c6c
Trichodesmium erythraeum c6c
Crocosphaera watsonii c6c 2
Crocosphaera watsonii c6c
Cyanothece c6c
Nostoc 7120 c6c
Anabaena variabilis c6c
Nostoc punctiforme c6c
Nodularia spumigena c6c
c_{6C}
Synechococcus 7942 c6
Bumilleriopsis filiformis c6
Porphyra yezoensis c6
Synechocystis c6
Pediastrum boryanum c6
Synechococcus 7002 c6
Acaryochloris marina c6
Arthrospira maxima c6
Tricondyloides elongatus c6
Chlamydomonas reinhardtii c6
Monoraphidium braunii c6
Scenedesmus obliquus c6
Nostoc 7120 c6
Nodularia spumigena c6
Aphanizomenon flos-aquae c6
c_6
Nostoc 7120 c550
Tricondyloides elongatus c550
Synechococcus c550
Synechocystis c550
c_{550}
Chlorobium tepidum 2 c550
Chlorobium limicola 2 c550
Chlorobium tepidum c550
Chlorobium limicola c550
c_{555}

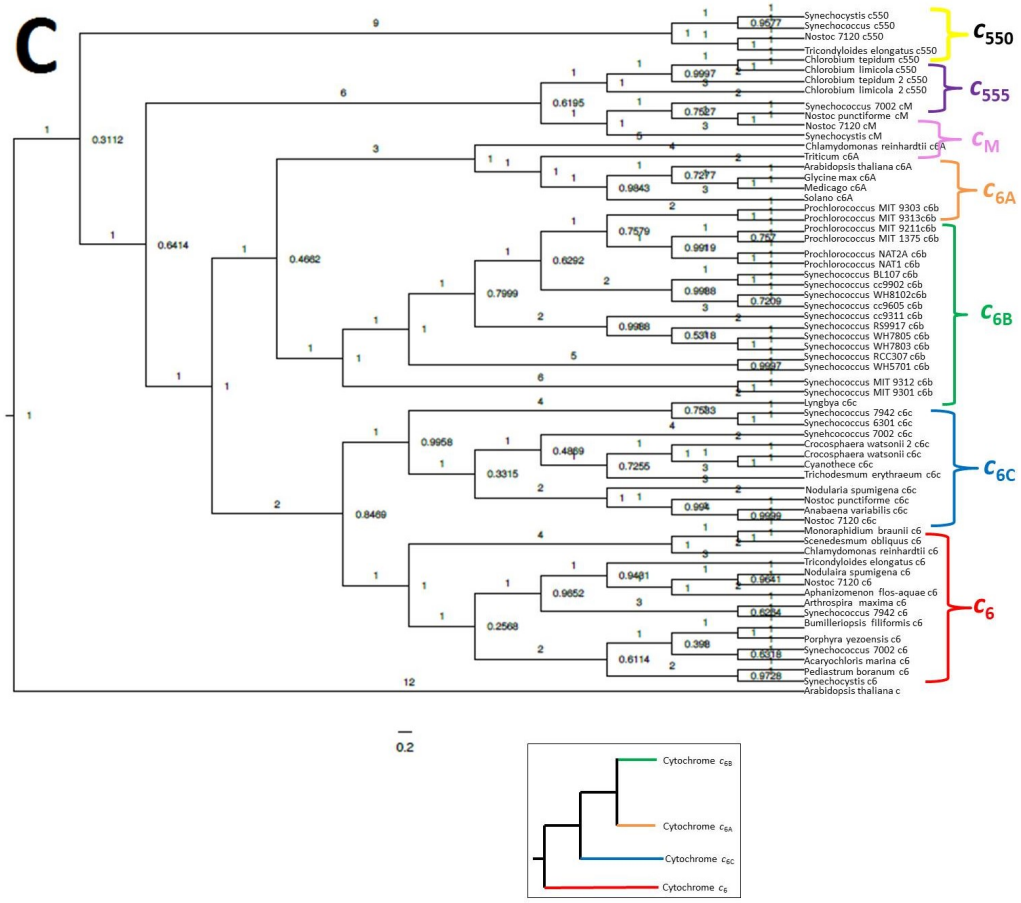


Figure 3.8: Phylogenetic trees built using the Maximum Likelihood (A), Parsimony (B) and Bayesian Inference (C) algorithms on the Bialek-ClustalX alignment. The original clusters from the Bialek et al. (2008) paper have been marked in the same colour brackets as seen in figures 3.1 and 3.6. Bootstrap values given at branchpoints. In the boxes are simplified forms of each tree demonstrating the phylogeny of only the c_6-type cytochromes, with cytochrome c_6 as the root.

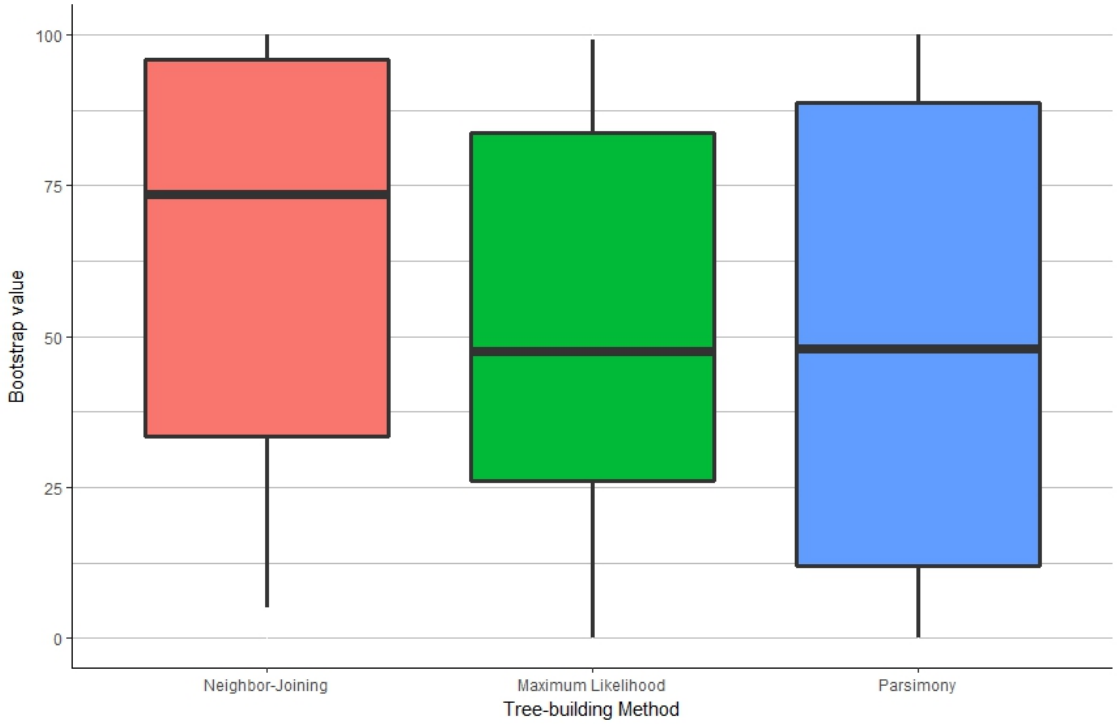


Figure 3.9: Distribution of the bootstrap values from three of the tree-building methods performed in MEGA7 on the alignment sourced from Bialek et al., (2008).

3.2.1.2. Creating a new phylogenetic tree using a wider range of species

3.2.1.2.1 Searching for cytochrome c_6 family proteins across photosynthetic organisms

In order to clarify the model of cytochrome c_6 evolution, genomic, transcriptomic and proteomic databases of a wide range of photosynthetic organisms were searched for c_6 family genes and proteins respectively. Once a putative gene, transcript or protein was identified from a BLAST search, the following criteria were used to confirm and annotate the sequence:

1. The sequence must contain a haem-binding motif (CXXCH) to be considered a type-I cytochrome (Barker and Ferguson, 1999).
2. When aligned with PETJ of *A. thaliana* and CYC4 of *C. reinhardtii,* if the sequence contained a glutamine at position 52 and 51 in PETJ and CYC4 respectively, then the sequence was labelled cytochrome c_6 as this would suggest a higher redox midpoint potential (Worrall *et al.*, 2007). Should the residue be hydrophobic instead (i.e. valine, isoleucine or leucine) the cytochrome would be likely to have a lower redox midpoint potential and be cytochrome $c_{6A/B/C}$.

3. In the same alignment with *A. thaliana* and *C. reinhardtii* c_{6A} described in point 2, if a low redox midpoint cytochrome had an insertion containing two cysteines, then the cytochrome was considered c_{6A}. If there was no insertion, the cytochrome was labelled $c_{6B/C}$.

As the specific difference(s) between cytochrome c_{6B} and c_{6C} are uncertain and based on phylogeny, the two were considered together. As well as searching amongst cyanobacteria, higher plants and eukaryotic green algae, the following photosynthetic organisms were considered:

- Haptophtes
- Glaucophytes
- Rhodophytes
- Ochrophyta
- Euglenids
- Cryptomonads
- Dinoflagellates
- Photosynthetic Rhizaria
- Purple phototrophic bacteria
- Green Sulphur bacteria

Within these clades, *c* type cytochromes were found in the species in a variety of combinations (figure 3.10). Cytochrome c_{6A}, initially believed to be exclusively found in the green plants and eukaryotic green algae, was also found in *Symbiodinium microadriaticum* and *Amphidinium carterae*, Dinoflagellate species thought to have obtained their plastid through secondary endosymbiosis of a rhodophyte (Dorrell *et al.*, 2015). Cytochromes $c_{6B/C}$ were found in cyanobacteria, ochrophytes, cryptomonads, photosynthetic rhizaria, haptophytes and rhodophytes, and cytochrome c_6 was found in cyanobacteria, Haptophytes, Glaucophytes, Rhodophytes, Eukaryotic green algae and Euglena. This is the first report of cytochrome $c_{6B/C}$ in Rhodophyta (figure 3.3). Interestingly, EST searches revealed many potential cytochrome $c_{6B/C}$ sequences in Delta and Gamma proteobacteria. Figure 3.11 shows an alignment demonstrating how a representative sequence from each clade has been identified as cytochrome c_{6A} or $c_{6B/C}$.

Clade	Contains cytochrome c_6	Contains cytochrome c_{6A}	Contains cytochrome $c_{6B/C}$
Embryophyta	X	✓	X
Chlorophyta	✓	✓	X
Haptophyta	✓	X	✓
Glaucophyta	✓	X	X
Rhodophyta	✓	X	✓
Ochrophyta	✓	X	✓
Euglenids	✓	X	X
Cryptomonads	✓	X	✓
Rhizaria	✓	X	✓
Dinoflagellata	✓	✓	✓
Cyanobacteria	✓	X	✓
Purple phototropic bacteria	✓	X	✓*
Green Sulphur bacteria	X	X	X

*Figure 3.10: Table of which photosynthetic clades contained an identified cytochrome c_6 family gene, transcript or protein. Ticks indicate that a potential gene, transcript (from EST searches) or protein matching that cytochrome has been found in at least one organism in that clade. *Cytochrome $c_{6B/C}$ homologues found in purple phototropic bacteria were found in Delta and Gamma proteobacteria.*

Examples of *c*-type cytochromes from these clades were sourced and the sequences included in the phylogenetic tree built in the subsequent section.

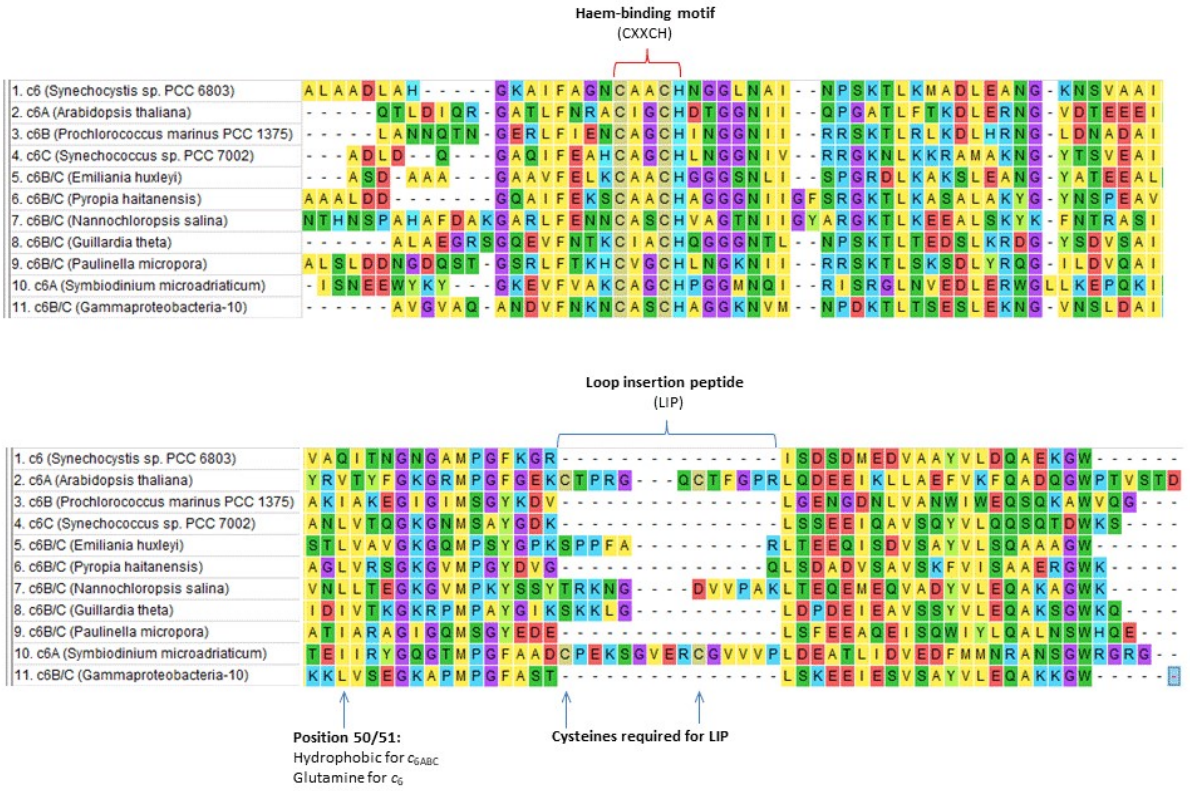


Figure 3.11: Alignment of cytochromes c_6, c_{6A}, c_{6B} and c_{6C} against examples of new potential c-type cytochromes found in photosynthetic organisms. Potential cytochrome c_{6A} in dinoflagellates sourced from Symbiodinium microadriaticum, and potential cytochromes $c_{6B/C}$ sourced from Emiliania huxleyi (haptophyte), Pyropia haitanensis (rhodophyte), Nannochloropsis salina (ochrophyte), Guillardia theta (cryptomonad), Paulinella micropora (Rhizaria) and Gammaproteobacter-10 (purple phototrophic bacteria). The haem-binding motif, position 50/51 and the LIP are pointed out to demonstrate how each sequence was estimated to belong to each c-type cytochrome.

3.2.1.2.2 Pre-existing phylogenetic trees were used to select peptide sequences for the wider alignment.

A possible weakness of the Bialek *et al.* (2008) analysis was the limited range of species used in the alignment. This is especially apparent in the cytochrome c_{6B} cluster, which only contains sequences from 2 genera: *Prochlorococcus* and *Synechococcus.* Therefore, a new alignment was created using sequences sourced from a wider range of photosynthetic organisms, including prokaryotes and eukaryotes. An ML tree previously created using an alignment of 756 protein sequences sampled across 65 cyanobacterial species was used to select prokaryotic organisms to use in the new alignment (Schirrmeister *et al.*, 2015). The tree was split into 6 major clusters (figure 3.12), with

clade 1 consisting of species yet to have annotated genomes published. Peptide sequences of cytochromes c_6 and $c_{6B/C}$ (thus named as they are yet to be assigned) were taken from a wide selection of organisms across this tree for the alignment study in order to get as wider coverage of cyanobacteria. Some sequences of cytochrome c_{6B} were also taken from the Bialek *et al.* (2008) alignment in order to see if the clustering still occurred. Cytochromes *c*, *c*-555, *c*-550 and c_M were excluded from the alignment as the aim of this study was to analyse the evolution and phylogeny of the c_6 family only.

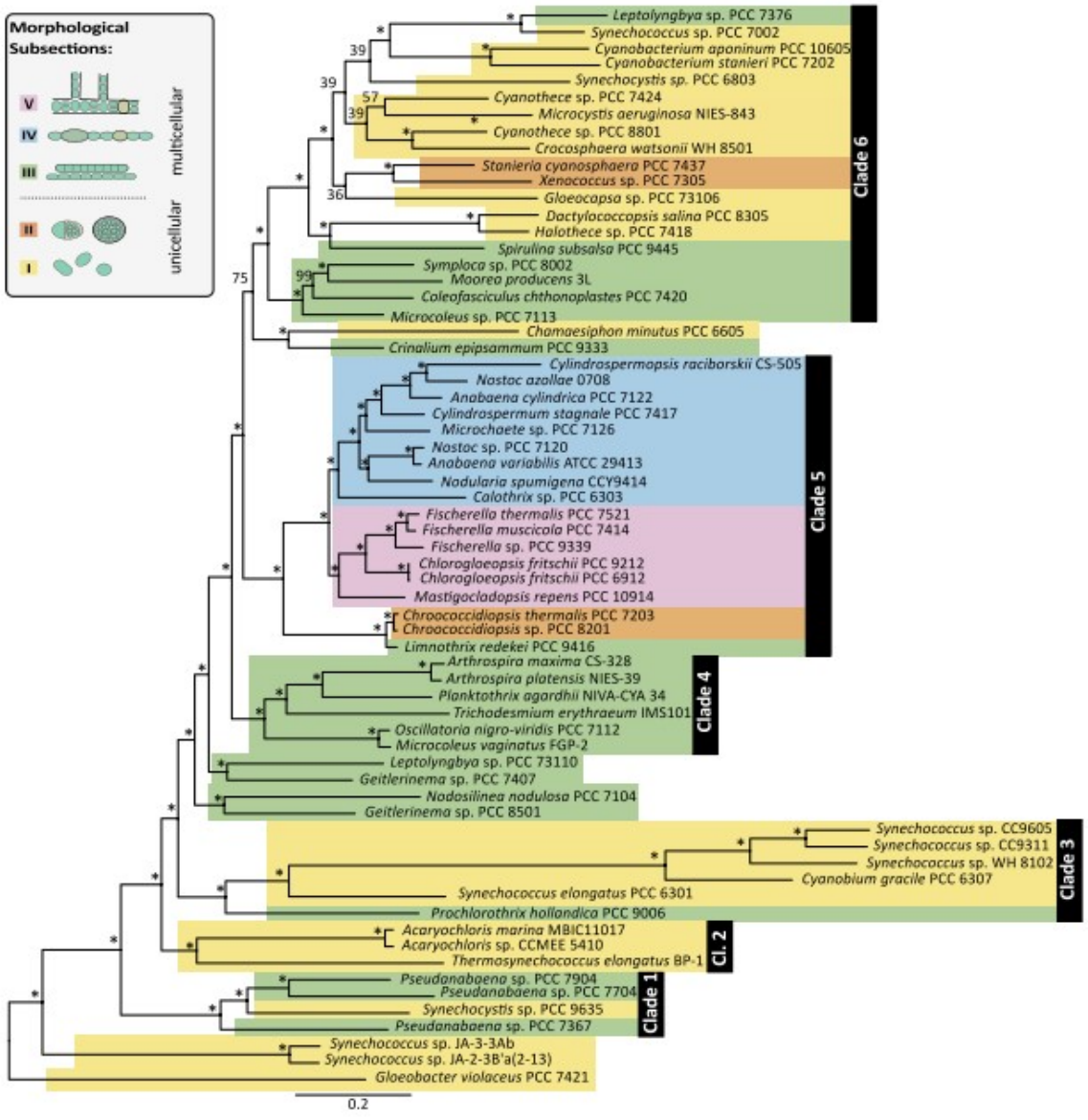


Figure 3.12: Maximum likelihood phylogenetic tree built using an alignment of 756 protein sequences across 65 cyanobacterial species. Colours indicate prior taxonomic groups identified through morphology, with a visual description of each morphology given in the legend. Bootstrap values of 100 are labelled with an asterisk. Figure taken from Schirrmeister et al., 2015.

For higher plants and eukaryotic green algae species, the phylogenetic tree built for Phytozome 12 (https://phytozome.jgi.doe.gov/pz/portal.html) was used (Goodstein *et al.*, 2012; figure 3.13). The tree was built using NCBI taxonomy and refined using reviews of relevant literature and expert input (Phillips, J. at Phytozome, personal communication). As with the cyanobacterial tree, examples of each clade were selected and sequences of cytochrome c_{6A} were obtained and added to the alignment. In the case of the eukaryotic green algae, sequences for cytochrome c_6 were also obtained.

Figure 3.13: Phylogenetic tree of higher plants and green eukaryotic algae taken from Phytozome 12. Tree was built based on data at https://phytozome.jgi.doe.gov/pz/portal.html# (date accessed 13/01/2019).

Finally, cytochrome c_6 family sequences from other photosynthetic organisms were added to the alignment. Cytochromes c_6, $c_{6B/C}$ and c_{6A} (where relevant) were taken from species in the haptophyte, glaucophyte, rhodophyte and dinoflagellate clades, as described in figure 3.10.

3.2.1.2.3 Using Muscle and Maximum Likelihood to build a phylogenetic tree using the wide range of cytochrome sequences

The new broader collection of peptide sequences were then aligned for building a more diverse phylogenetic tree. The signal peptides of each sequence were identified (through alignments with known coding regions of cytochrome c_6) and removed before the peptides were aligned using the Muscle algorithm. The OMD was then calculated for the alignment (figure 3.5) and found to be lower than for the previous alignments, and close to 1. A Maximum Likelihood phylogenetic tree was built using the gamma-distributed WAG model plus invariable sites in MEGA7 (figure 3.14).

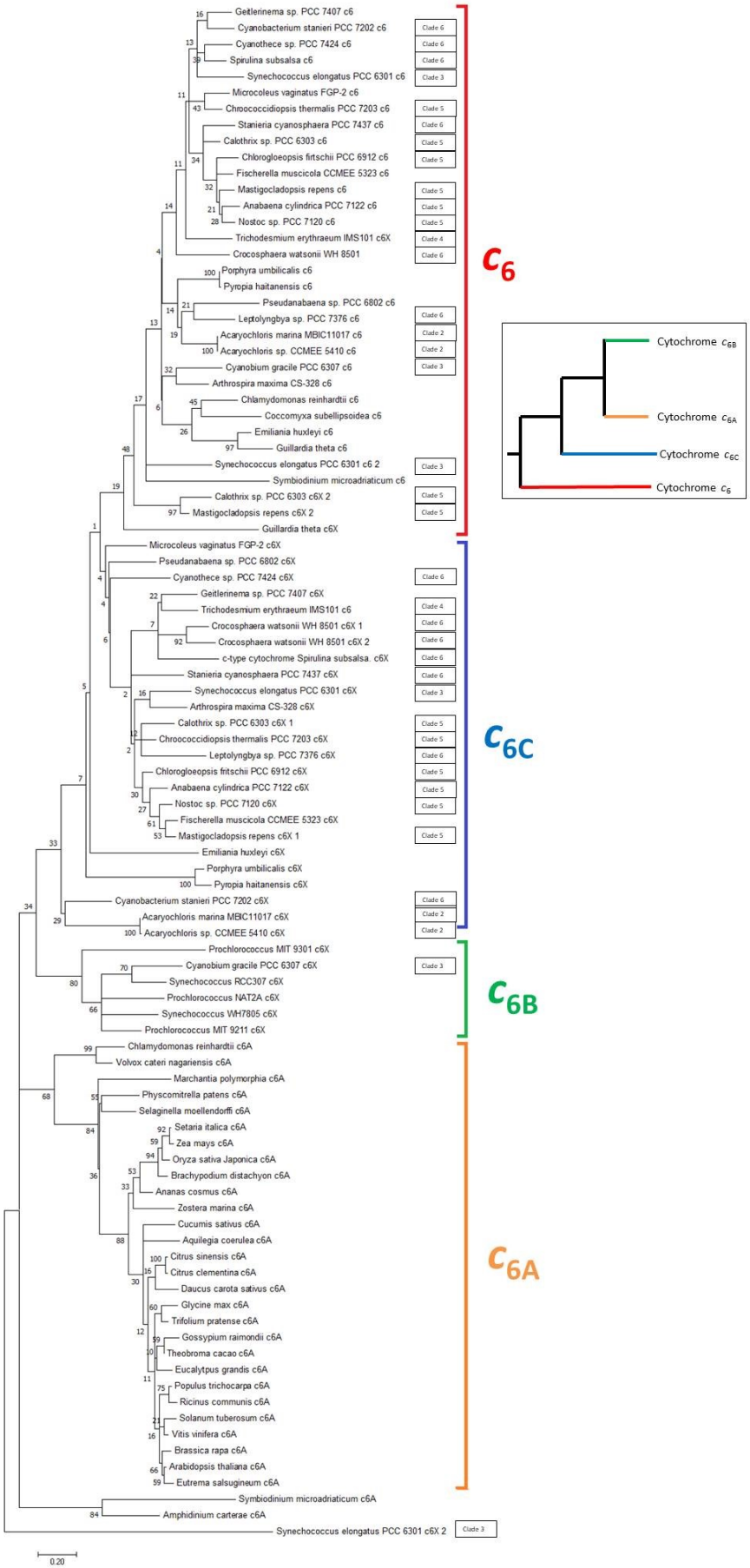

Geitlerinema sp. PCC 7407 c6
Cyanobacterium stanieri PCC 7202 c6
Cyanothece sp. PCC 7424 c6
Spirulina subsalsa c6
Synechococcus elongatus PCC 6301 c6
Microcoleus vaginatus FGP-2 c6
Chroococcidiopsis thermalis PCC 7203 c6
Stanieria cyanosphaera PCC 7437 c6
Calothrix sp. PCC 6303 c6
Chlorogloeopsis fritschii PCC 6912 c6
Fischerella muscicola CCMEE 5323 c6
Mastigocladopsis repens c6
Anabaena cylindrica PCC 7122 c6
Nostoc sp. PCC 7120 c6
Trichodesmium erythraeum IMS101 c6X
Crocosphaera watsonii WH 8501
Porphyra umbilicalis c6
Pyropia haitanensis c6
Pseudanabaena sp. PCC 6802 c6
Leptolyngbya sp. PCC 7376 c6
Acaryochloris marina MBIC11017 c6
Acaryochloris sp. CCMEE 5410 c6
Cyanobium gracile PCC 6307 c6
Arthrospira maxima CS-328 c6
Chlamydomonas reinhardtii c6
Coccomyxa subellipsoidea c6
Emiliania huxleyi c6
Guillardia theta c6
Synechococcus elongatus PCC 6301 c6 2
Symbiodinium microadriaticum c6
Calothrix sp. PCC 6303 c6X 2
Mastigocladopsis repens c6X 2
Guillardia theta c6X
Microcoleus vaginatus FGP-2 c6X
Pseudanabaena sp. PCC 6802 c6X
Cyanothece sp. PCC 7424 c6X
Geitlerinema sp. PCC 7407 c6X
Trichodesmium erythraeum IMS101 c6
Crocosphaera watsonii WH 8501 c6X 1
Crocosphaera watsonii WH 8501 c6X 2
c-type cytochrome Spirulina subsalsa. c6X
Stanieria cyanosphaera PCC 7437 c6X
Synechococcus elongatus PCC 6301 c6X
Arthrospira maxima CS-328 c6X
Calothrix sp. PCC 6303 c6X 1
Chroococcidiopsis thermalis PCC 7203 c6X
Leptolyngbya sp. PCC 7376 c6X
Chlorogloeopsis fritschii PCC 6912 c6X
Anabaena cylindrica PCC 7122 c6X
Nostoc sp. PCC 7120 c6X
Fischerella muscicola CCMEE 5323 c6X
Mastigocladopsis repens c6X 1
Emiliania huxleyi c6X
Porphyra umbilicalis c6X
Pyropia haitanensis c6X
Cyanobacterium stanieri PCC 7202 c6X
Acaryochloris marina MBIC11017 c6X
Acaryochloris sp. CCMEE 5410 c6X
Prochlorococcus MIT 9301 c6X
Cyanobium gracile PCC 6307 c6X
Synechococcus RCC307 c6X
Prochlorococcus NAT2A c6X
Synechococcus WH7805 c6X
Prochlorococcus MIT 9211 c6X
Chlamydomonas reinhardtii c6A
Volvox cateri nagariensis c6A
Marchantia polymorphia c6A
Physcomitrella patens c6A
Selaginella moellendorffi c6A
Setaria italica c6A
Zea mays c6A
Oryza sativa Japonica c6A
Brachypodium distachyon c6A
Ananas cosmus c6A
Zostera marina c6A
Cucumis sativus c6A
Aquilegia coerulea c6A
Citrus sinensis c6A
Citrus clementina c6A
Daucus carota sativus c6A
Glycine max c6A
Trifolium pratense c6A
Gossypium raimondii c6A
Theobroma cacao c6A
Eucalytpus grandis c6A
Populus trichocarpa c6A
Ricinus communis c6A
Solanum tuberosum c6A
Vitis vinifera c6A
Brassica rapa c6A
Arabidopsis thaliana c6A
Eutrema salsugineum c6A
Symbiodinium microadriaticum c6A
Amphidinium carterae c6A
Synechococcus elongatus PCC 6301 c6X 2
Clade 2
Clade 3
Clade 4
Clade 5
Clade 6
0.20
c_6
c_{6C}
c_{6B}
c_{6A}
Cytochrome c_{6B}
Cytochrome c_{6A}
Cytochrome c_{6C}
Cytochrome c_6

Figure 3.14: Maximum likelihood phylogenetic tree built from the alignment of 95 peptide sequences of cytochrome c_6 family proteins from a wide range of photosynthetic organisms. Peptides are tentatively coloured by the cytochrome subfamily they belong to. Cytochrome c_{6X} refers to cytochromes that could be c_{6B} or $_{c6C}$, based on the presence of a CXXCH haem-binding motif and a hydrophobic residue at the equivalent of position 51 in C. reinhardtii numbering. "Clade" boxes beside cyanobacteria sequences denote which clade they were assigned in figure 3.12. Alignment created using Muscle algorithm. Bootstrap values given at branchpoints. In the box to the right is a simplified tree demonstrating the phylogeny of only the c_6-type cytochromes, with cytochrome c_6 as the root.

The average of the bootstrap values of the tree was again lower than that of the Bialek *et al.* (2008) tree (figure 3.15). The subfamilies of cytochrome c_6 do appear to cluster with c_{6A}. The dinoflagellate cytochrome c_{6A} sequences diverge at the base of the c_{6A} cluster, but this may be a consequence of the long edges leading to those sequences. This may also explain the placing of the, putative $c_{6B/C}$ protein sequence from *S. elongatus* PCC 6301. The *Trichodesmium erythraeum* cytochromes c_6 and $c_{6B/C}$ appear in the opposite clusters from expected. This may suggest that the peptide sequence currently annotated c_6 in NCBI is in fact the $c_{6B/C}$ and vice versa.

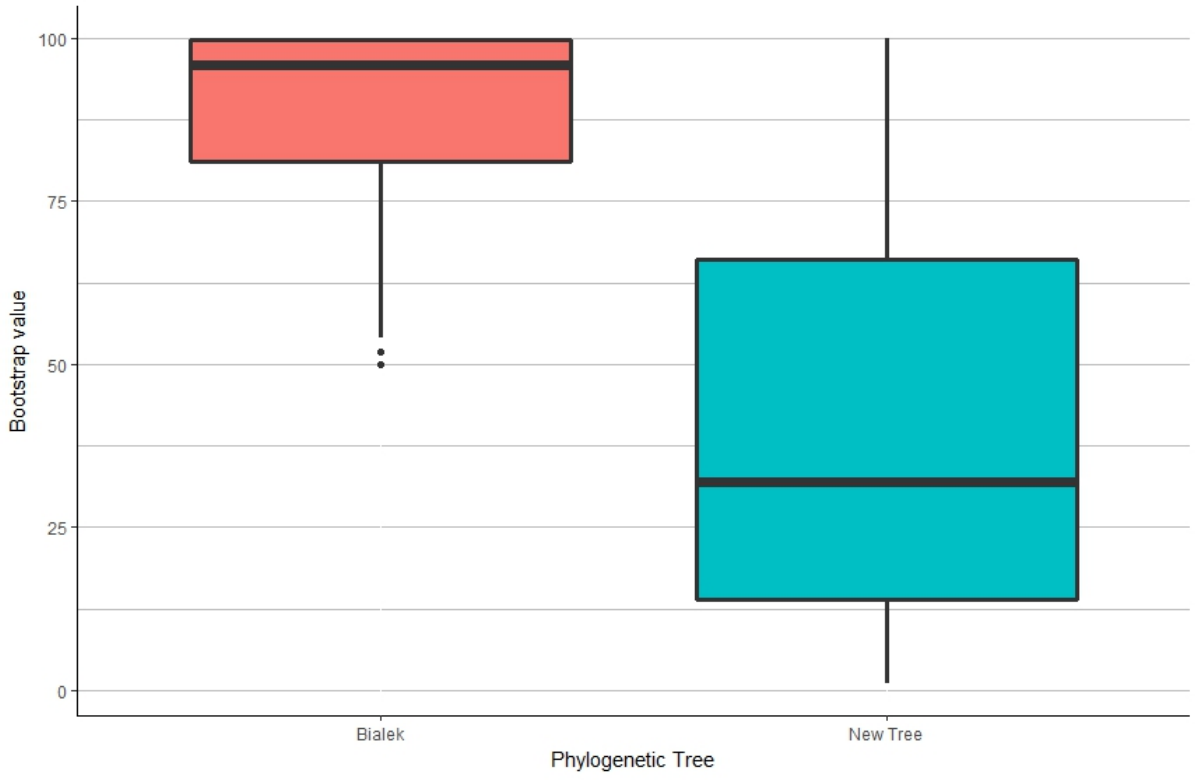


Figure 3.15: Comparison of the distribution of bootstrap values between the phylogenetic tree built in Bialek et al, 2008 and the phylogenetic tree built using a wider range of sequences in figure 3.14 ("New Tree").

The sequences labelled cytochrome c_{6B} and c_{6C} in Bialek *et al.* (2008) separated into two clusters in figure 3.14. If the root of the tree were on the edge leading to cytochrome c_6, it would suggest that cytochromes c_{6A} and c_{6B} share a more common recent ancestor than either does with cytochrome c_{6C}. It is also worth noting that the bootstrap values between all of these clusters were very low, thus reducing the reliability of the segregation of said clusters.

The distribution of sequences in the c_{6A} cluster appeared to match the distribution of organisms in the Phytozome tree. Setting aside the highly divergent dinoflagellate sequences, the first branching of the c_{6A} cluster (rooting it using the other sequences in the tree) separated embryophytes from the chlorophytes, the next branch in the embryophytes led to the angiosperms, and the subsequent branch from that led to the eudicots. Within the Eudicots the distribution was less well matched to the Phytozome tree, although all the Brassicaeae appeared to cluster together. The distributions of the cytochrome c_6 and $c_{6B/C}$ clusters did appear to match those of the cyanobacteria tree from figure 3.12 to some extent, with the clades 5 and 6 largely clustering together with low bootstrap support, particularly in the c_{6C} cluster, and clade 2 clustering with a high bootstrap level. Clades 3 and 4 were the exception to this, with no clear clustering observed.

The sequences included from non-green photosynthetic eukaryotes appeared to share a most recent common ancestor to cytochromes c_6 and c_{6C}. The Rhodophyta and the haptophyte *Emiliana huxleyi* $c_{6B/C}$ share a common ancestor most recently with cytochrome c_{6C}. The cryptophyte *Guillardia theta* $c_{6B/C}$ sequence shares a more recent common ancestor to cytochrome c_6 than $_{c6B/C}$, although given the position of the *Guillardia theta* sequence, this depends on where the root is placed in relation to the cytochrome c_6 sequences as if the root was placed between the *G. theta* sequence and the c_6 family, the sequence would cluster with one of the other subfamiles. Therefore combined with the lack of a LIP-style insertion, the rhodophyte and haptophyte sequences were labelled cytochrome $c_{6B/C}$ as opposed to c_{6A}.

3.2.1.3. A new proposed model of cytochrome c_6 family ancestry and evolution

The search for c_6 family cytochromes in a wide range of photosynthetic organisms has led to an updated model of the evolution of this protein family (figure 3.16). Firstly, the original model did not separate cytochrome c_{6A} from $c_{6B/C}$, therefore in the updated model the two subfamilies of protein

are treated separately. Another important update to the model is the inclusion of cytochrome $c_{6B/C}$ in the red algae, countering the idea from figure 3.3 that cytochrome $c_{6B/C}$ was lost before Rhodophyta evolved. The model has also separated glaucophytes from red algae as it appears that the former has lost cytochrome $c_{6B/C}$, and the dinoflagellates that gained a plastid from rhodophytes have been added.

The proposed model begins with the ancestor to the chloroplast containing the genes for plastocyanin and cytochromes c_6 and the low potential cytochromes $c_{6B/C}$. The latter are likely to have appeared through duplication of cytochromes c_6. This situation is observed in most extant cyanobacteria. After primary endosymbiosis, cytochrome c_6 was retained in all subsequent clades analysed except green plants, and plastocyanin was lost in the red algae, glaucophytes and any organism that obtained them as a symbiont through secondary endosymbiosis. Before the divergence of green plants and green algae, cytochrome $c_{6B/C}$ gained an insertion (the LIP) and became cytochrome c_{6A}. This was lost in euglenids either before or after the secondary endosymbiosis that gave rise to them. Cytochrome $c_{6B/C}$ was also lost in the glaucophytes. The origin of the cytochrome c_{6A} in the dinoflagellates is unclear, given the lack of cytochrome c_{6A} in red algae, and is discussed further below.

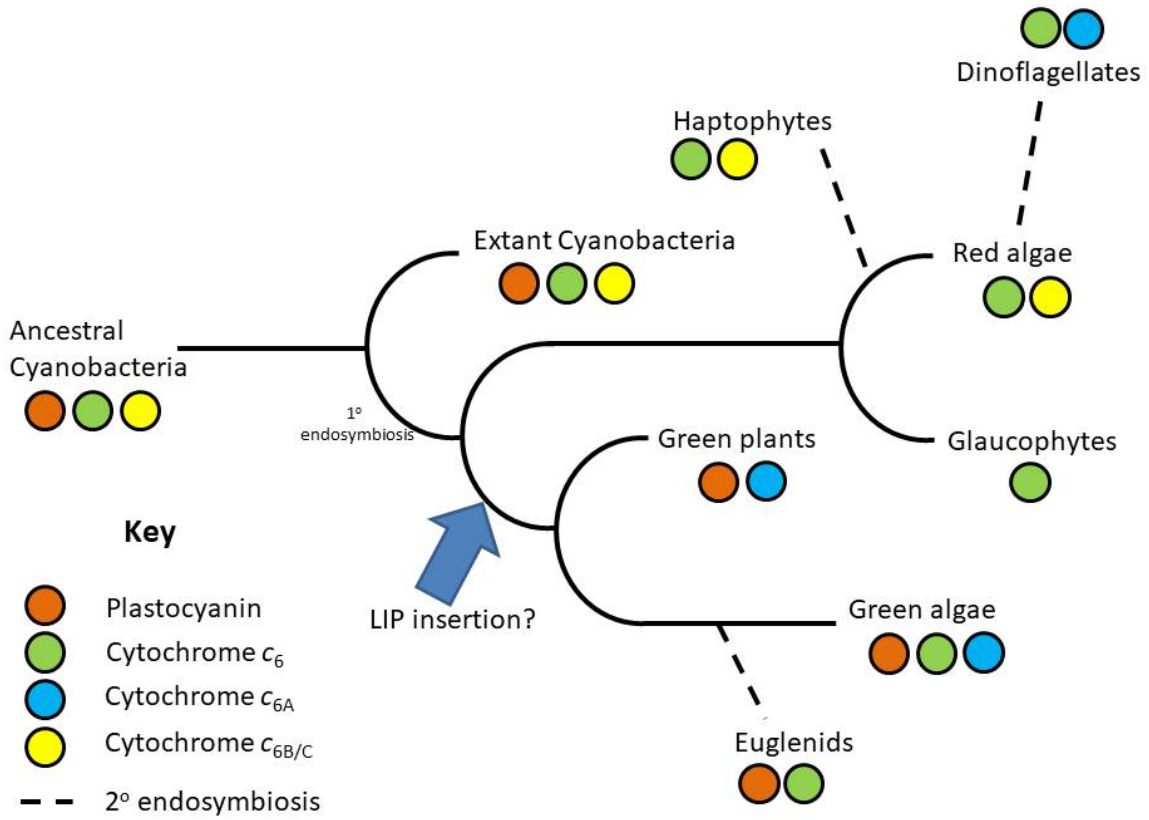


Figure 3.16: Updated ancestry model for the cytochrome c_6 family in photosynthetic organisms. Adapted from figure 3.3 and Howe et al., 2015. In this context, 'Dinoflagellates' refers to the peridinin-containing lineage.

3.2.2. Analysing the conservation of the potential linkage between cytochrome c_{6A} and trypsin family protease and Glutathione-S-transferase hypothetical proteins

In *A. thaliana* the coding region for cytochrome c_{6A} (AT5g45040) has genes for a trypsin-family protease (*TFP*, AT5g45030) and a glutathione-S-transferase (*GST*, AT5g45020) located within 10 kbp, with the *TFP* region overlapping with c_{6A}. This close linkage was searched for across the Viridiplantae genomic databases uploaded onto Phytozome 12.1 (https://phytozome.jgi.doe.gov/pz/portal.html#). For each organism in Phytozome, the homologue of *petJ*/AT5g45040 was identified along with the closest homologues for *TFP* and *GST*. It was then noted for either gene if it was located within 10kbp of *petJ* (close linkage), on the same chromosome as *petJ* but not within 10 kbp (linkage) or located on a separate chromosome from *petJ* (no linkage). It was also noted when no homologue could be found for any of the three genes involved (no data). This information was then mapped on the Phytozome phylogenetic tree for Viridiplantae (figure 3.17).

Figure 3.17: Phylogenetic trees of Viridiplantae colour-coded to represent the level of linkage between petJ and TFP (A) or GST (B). Organisms shown in green have close linkage (within 10 kbp),

organisms in yellow have linkage (same chromosome), organisms in red have no linkage and organisms in black had insufficient data to come to a conclusion. Exact distances between genes for each organism can be found in the supplementary table S2.

In the Brassicaceae clade, the linkage between *petJ* and the respective hypothetical genes is present in many species, and frequently close. However, for organisms more distantly related to the Brassicaceae the linkage is more distant and less likely to exist at all. This suggests that the linkage between both *TFP* and *GST* and *petJ* is incidental and occurred by chance in Brassicaceae. This is supported by observation of high synteny between *A. thaliana* and other members of the Brassicaceae family (Schranz et al., 2007). Assuming that cytochrome c_{6A} does not play a separate role in *Brassica* relative to the remaining Viridiplantae, there is no certain link between function and the linkage of these genes. Furthermore, coexpression studies performed in *A. thaliana* and *C. reinhardtii* (Gioirgi *et al.*, 2013; Romero-Campero *et al.*, 2016) do not show any other genes to be consistently coexpressed with cytochrome $_{c6A}$, including TFP and GST.

3.3. Discussion

3.3.1. The new proposed model of cytochrome c_6 family ancestry and evolution

After endosymbiosis plastid cytochrome $c_{6B/C}$ appears to have followed two different paths. One such path was to remain as cytochrome $c_{6B/C}$ in the Rhodophyta, and Haptophytes following the secondary endosymbiosis. The other path was to gain an insertion containing two cysteine residues (the LIP) and become cytochrome c_{6A}. This latter path then led to the green eukaryotic algae and higher plants. In higher plants cytochrome c_6 was lost, with plastocyanin remaining as the sole electron transporter between cytochrome b_6f and PSI. In eukaryotic algae both cytochrome c_6 and plastocyanin remained, with cytochrome c_{6A} being lost after the secondary endosymbiosis into the *Euglena* lineage. The fact that no green alga or higher plant contains cytochrome $c_{6B/C}$, and no Haptophyte, Cyanobacterium, Rhodophyte or Glaucophyte contains a cytochrome c_{6A} implies that the LIP insertion occurred in the $c_{6B/C}$ gene in the green algal lineage very shortly before or after its divergence from the rhodophytes and glaucophytes.

3.3.1.1. Further refinements to make in the model of cytochrome c_6 family ancestry

There are further refinements that are needed in the model proposed in figure 3.16. One thing that is unclear is whether cytochrome c_6 or $c_{6B/C}$ came first and which appeared as a result of duplication

and residue substitution. As suggested before, the duplication event of cytochrome c_6 to form a separate lineage in which subsequent substitution events occurred that led to a lower redox midpoint potential cytochrome is still most likely to have occurred in the cyanobacterial ancestor to the chloroplast. The essential nature of cytochrome c_6 as a core component of the photosynthetic electron transport chain may suggest that it came first, given that plastocyanin is believed to have evolved as a photosynthetic electron carrier after cytochrome c_6 (Howe *et al.*, 2005). The presence of both cytochrome c_6 and $c_{6B/C}$ in extant cyanobacteria provides evidence that this duplication and substitution event occurred prior to endosymbiosis, assuming that the lineages containing both genes diverged before the endosymbiotic event.

Increasing the scope of the searches behind the model would also improve its accuracy. Despite searching genome, transcriptome and proteome data, many of the photosynthetic lineages including Rhodophyta, cryptomonad, Ochrophyta, Glaucophyta and Haptophyta, had fewer databases available than for cyanobacteria and higher plants. As more organisms are sequenced and annotated, they can be searched for cytochrome c_6 proteins and the model can be further renewed.

The inclusion of the dinoflagellates in the model has brought new questions, as the origin of cytochrome c_6 family proteins appears more complicated in this lineage. The organisms that contained a cytochrome c_{6A}-like protein (*Amphidinium carterae* and *Symbiodinium adriaticum*) have a peridinin-containing plastid, which was gained through secondary acquisition of a rhodophyte plastid (Keeling 2010; Dorrell *et al.*, 2015). The rhodophytes were shown in this study to contain a cytochrome $c_{6B/C}$. Therefore the question remains as to whether the dinoflagellate cytochrome c_{6A} is functionally the same as the green eukaryotic lineage cytochrome c_{6A}. The peptides do have the amino acid substitution suggesting a lower redox midpoint potential (figure 3.11), and contain an insertion that is in the same location as the LIP and contains two cysteine residues. The phylogenetic analysis did show that the c_{6A}-style sequences clustered with other c_{6A} sequences, although their position must be viewed with caution, given the length of the branches on which they were located (figure 3.14). Thus it appears that dinoflagellates may contain a c_{6A} gene, and therefore the question of the appearance of a LIP-containing cytochrome c_6 in dinoflagellates becomes: did the insertion happen again in a $c_{6B/C}$ peptide from the rhodophyte, or was lateral gene transfer from a green eukaryote involved? There has been evidence of lateral gene transfer to dinoflagellates from a variety of organisms (Takishita *et al.*, 2003; Hackett *et al.*, 2005; Chan *et al.*, 2012; Wisecaver *et al.*, 2013). Again, further bioinformatic analysis of cytochrome c_{6A} in dinoflagellates as more genomes are published should be combined with functional molecular genetic techniques to determine the nature of this c_{6A}-like peptide.

An interesting discovery of this study was the presence of potential cytochrome $c_{6B/C}$ sequences in the Deltaproteobacteria and Gammaproteobactera. The *c*-type sequences found did not have a LIP-like insertion, but did have a hydrophobic amino acid align with position 51 in *C. reinhardtii* cytochrome c_{6A}, suggesting that the cytochrome found may be of low redox potential. The protein appearance in these lineages could either be from horizontal gene transfer, or from independent evolution of a *c*-type cytochrome, as the sequence was not found in all proteobacteria. The Deltaproteobacteria and Gammaproteobacteria clades vary in physiology, with some being photosynthetic, chemoautotrophic or neither (Garrity *et al*., 2006), so further investigation into the differences between species that contain a $c_{6B/C}$ and those that do not needs to be performed before this clade can be included in the current model of c_6 family ancestry.

3.3.2. Comparison of phylogenetic trees and implications for cytochromes c_{6B} and c_{6C}

Further investigation into the phylogenetics and ancestry of cytochromes c_{6B} and c_{6C} provides evidence that the two genes are orthologs that are functionally the same. Firstly, the sequences that contributed to the cytochrome c_{6B} clade in the tree built by Bialek *et al.* (2008) all came from two genera of cyanobacteria: *Synechococcus* and *Prochlorococcus*. This means that these sequences would be likely to cluster to the exclusion of the other sequences, which, apart from some *Synechococcus* species were from differing sources. This may also explain why the cytochrome c_{6B} sequences used in figure 3.14 clustered together again, rather than because they encode a paralogue with a different function. This is supported by a phylogenetic tree of cyanobacteria built using an alignment of conserved marker genes (Walter *et al.*, 2017), in which the *Synechococcus* and *Prochlorococcus* that were in the c_{6B} cluster grouped to the exclusion of the other cyanobacteria, which were represented by species found in the c_{6C} cluster in Bialek *et al.* (2008). Therefore it appears likely that the observed clusters of cytochrome c_{6B} and c_{6C} instead reflect the ancestry of the cyanobacteria in general, and thus cytochrome c_{6B} and c_{6C} could be functionally identical orthologs.

The idea that cytochrome c_{6B} and c_{6C} are functionally similar orthologs is supported by the observation that amongst these species used, none contained both a cytochrome c_{6B} and a c_{6C}, even amongst the *Synechococcus*. The presence of both genes in one organism would suggest that they perform differing functions. In the newer tree multiple cytochrome $c_{6B/C}$ sequences were used from the same species, but none of the sequence pairs clustered into the cytochrome c_{6B} and c_{6C} clusters.

The average bootstrap values observed with the trees inferred in this study were much lower than seen in Bialek *et al.* (2008). This is surprising, especially when one of the trees presented here used

the Bialek-ClustalX alignment used and Neighbor-Joining. As most of the sequences aligned were of around 50 amino acids in length it is unsurprising that low bootstrap values were obtained in this study. Unfortunately, it would not be possible to make the sequences longer and therefore the low bootstrap values should allow for a small element of scepticism to the data observed. However, it is worth noting that the software used to build the trees in this study (MEGA7) was not the same as the software used by Bialek *et al.* in 2008 (PHYLIP), therefore this difference in program may have contributed to the differing bootstrap measurements. For example, the amino acid substitution model was likely to have been different between the trees built in this study (which used the Jones-Taylor-Thornton model; Jones *et al.*, 1992) and those built by Bialek *et al* in 2008. Unfortunately the substitution model used in the latter study was not stated, but PHYLYP is only capable of running the Dayhoff PAM model or Kimura method, not the more modern Jones-Taylor-Thornton (Reteif, 2000).

Regardless, as it was also shown that the spread of the alignment was too large for an NJ tree, the method used to build the tree was arguably flawed. Given the short sequences used, it is not clear that any other method would be more reliable, however.

The final major difficulty found in this study is that the associations of the cytochrome c_{6B} and cytochrome c_{6C} clusters with the cytochrome c_{6A} and cytochrome c_6 clusters respectively were not reproducible in replicate trees. The model proposed by Bialek *et al.* (2008) was that cytochrome c_{6B} shared a more recent common ancestor with cytochrome c_{6A}, and cytochrome c_{6C} shared a more common ancestor with c_6 (figure 3.18A). However, the phylogenetic trees built in this study found that cytochromes c_{6A} c_{6B} and c_{6C} shared more recent common ancestry with each other than with cytochrome c_6 (figure 3.18B), using the same outgroups for rooting. The latter topology suggests that cytochrome c_{6A} could have formed as a derivative of $c_{6B/C}$ after the gene duplication of cytochrome c_6 and $c_{6B/C}$, which is consistent with the model shown in figure 3.16.

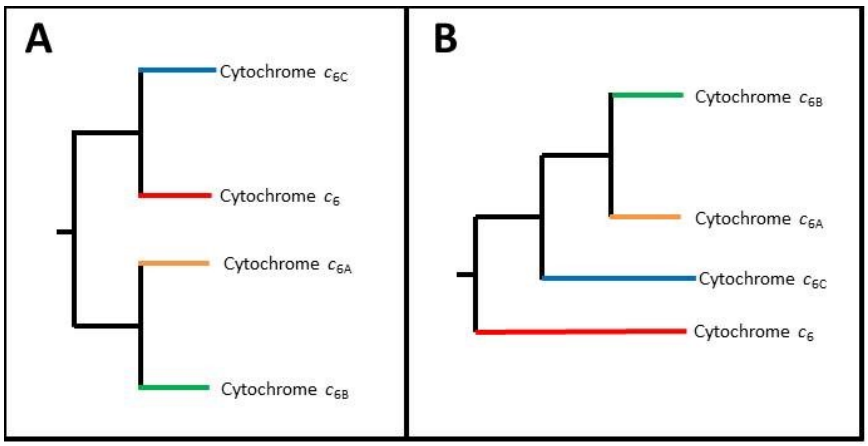


Figure 3.18: Comparison of the simplified phylogenetic trees proposed by Bialek et al., 2008 (A) and observed in a majority of the phylogenetic trees produced in this study (B). Branch lengths are not to scale.

The new tree built with only cytochrome c_6 family proteins (figure 3.14) had some advantages over Bialek *et al.* (2008), but also relies on some assumptions. One advantage is that the exclusion of less relevant *c*-type cytochromes such as c_M allowed only the relationship between the c_6-type cytochromes to be shown, and possibly contributed to the lower OMD observed. The phylogeny of the cyanobacterial sequences of the tree was generally consistent with the phylogenetic tree based on 756 protein sequences (Schirrmeister *et al.* (2015) and figure 3.12) with regard to clades 2, 5 and 6. However, clades 3 and 4 did not appear to cluster together, and the distribution of species did not exactly match that of Schirrmeister *et al*. (2015). This could be due to the short cytochrome sequence leading to errors in the topology. The mismatch in distribution was also observed with cytochrome c_6, a protein that should theoretically cluster in a similar pattern to figure 3.12. Therefore, the fact that the mismatch occurred with both sets of cytochrome c_6 subfamilies suggests that it may indeed be a consequence of using short sequences. The cytochrome c_{6A} cluster matched well with the Phytozome tree of figure 3.13. The clades of a higher taxonomic level in Viridiplantae clustered together, with the clustering not matching so well when considering lower taxonomic levels such as families. This could again be a result of short sequences leading to unreliable topology with closely related groups, as supported by the low bootstrap values. The cytochrome c_{6X} sequences of the rhodophytes, glaucophytes and haptophytes appeared to cluster more closely to $c_{6B/C}$ than c_{6A}, unlike those of the dinoflagellates. This agrees with the observation that there is no LIP in these sequences, and may suggest that the cytochrome c_{6A} sequences have been subjected to different selection patterns that the cytochrome $c_{6B/C}$ ones, which may be consistent with the presence of the LIP in the former but not the latter.

Therefore, there is insufficient evidence to suggest a separation of cytochrome c_{6B} and c_{6C} from a phylogenetic and evolutionary perspective. This should be confirmed through further molecular genetics and biochemical analysis of cytochrome c_{6B} to test if its function is similar to c_{6C}, as there have been functional experiments performed on the latter (Vasudevan, 2019) but not on the former. For example, analysing the growth phenotype of knockout mutants of c_{6B} under varying light intensities, as well as determining transcriptomic and proteomic changes in said knockouts in a similar fashion as has been done with cytochrome c_{6C}, would allow for functional comparisons between c_{6B} and c_{6C}. Determining the redox midpoint potential of cytochrome c_{6B} could also reveal

more about its function, as well as provide further evidence to assess the principle for predicting low redox cytochromes in this study (figure 3.11). If the functions of cytochromes c_{6B} and c_{6C} are the same, I propose that both genes are presented as being cytochrome c_{6B}.

3.3.3. Cytochrome c_{6A} has no known conserved linkage

This study did not find consistent linkage between either *TFP* or *GST* and the cytochrome c_{6A} gene. Widespread co-expression databases of *C. reinhardtii* and *A. thaliana* thus far have found no genes that co-express significantly with cytochrome c_{6A}, with any gene that could potentially have co-expression being found also to have a low base level expression. This low expression level of cytochrome c_{6A} will make co-expression studies particularly difficult, but should not be ignored.

4. Characterisation of cytochrome c_{6A} mutant lines through under stress conditions and chlorophyll fluorescence

4.1. Introduction

4.1.1. *Chlamydomonas* implements a photoprotective response to high and fluctuating light

Chlamydomonas reinhardtii must be able to survive in light conditions ranging from very high intensity to periods of darkness. This is due not only to the diurnal cycle, but also to changes in depth underwater in freshwater habitats and weather fluctuations over the course of a day. Periods of high light can be damaging to photosynthetic organisms as the increased flux of electrons through the electron transport chain can lead to a bottleneck effect as the rate of PSII water oxidation exceeds that of $NADP^+$ reduction and ATP synthesis (Papageorgiou, 2007). This in turn leads to the energy of the electrons transferring to molecular oxygen, generating reactive oxygenic species (ROS) which will indiscriminately attack and damage a wide range of important biomolecules (Halliwell, 1984). Therefore *C. reinhardtii* has evolved intricate photoprotective responses to survive varying light intensities (Erickson *et al.*, 2015), which has been reviewed in section 1.2.

4.1.2. Different aspects of photoprotection can be analysed using specific chemicals

4.1.2.1. Rose Bengal generates singlet oxygen when illuminated

Rose Bengal (RB, Figure 4.2) is a pink chemical often used for staining purposes (Feenstra *et al.*, 1992). In the presence of light, RB generates singlet oxygen (Neckers, 1989), and therefore is often used in studies to induce exogenous singlet oxygen stress, including in *C. reinhardtii* (Fischer *et al.*, 2004). In *C. reinhardtii*, RB has been used to 'prime' cells against singlet oxygen stress, with cultures that had been exposed to a small amount for 2 hours able to survive much higher concentrations in subsequent tests (Ledford *et al.*, 2007). This suggested that singlet oxygen exposure leads to a change in expression in the cells that help alleviate ROS stress. Analysis of *C. reinhardtii* mutants resistant to RB exposure led to the identification of the transcription factor SOR1 (Fischer *et al.*, 2012), and RB has been used to induce singlet oxygen stress in transcriptomic analysis of the stress response (Fischer *et al.*, 2005).

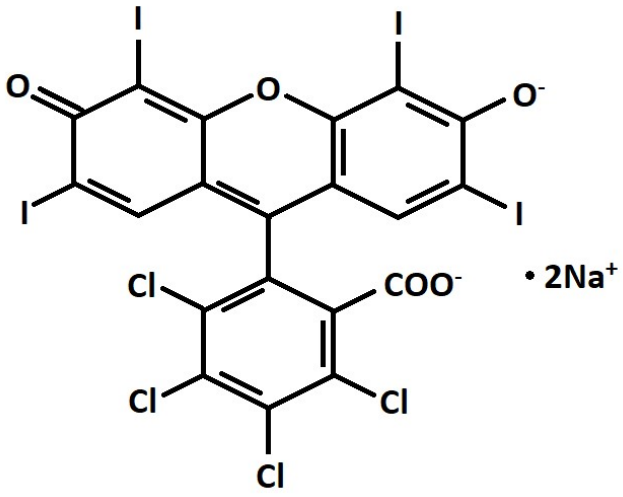


Figure 4.1: Chemical structure of Rose Bengal as a disodium salt.

4.1.2.2. The herbicide DBMIB can target PSII or cytochrome b_6f

Dibromothymoquinone (DBMIB) is a plastoquinone analogue that can competitively bind the quinone binding site (Q_0) on the cytochrome b_6f complex (Chain and Malkin, 1979; figure 4.3A). However, there is also evidence that DBMIB is able to quench excited chlorophyll in the LHC around PSII (Bukhov *et al.*, 2003), as well as studies that show DBMIB is able to oxidise Q_A^- at PSII (Belatik *et al.*, 2013). DBMIB therefore acts to decrease both plastoquinone reduction at PSII and plastoquinol binding to cytochrome b_6f, and can be used to interrupt both linear and cyclic electron flow.

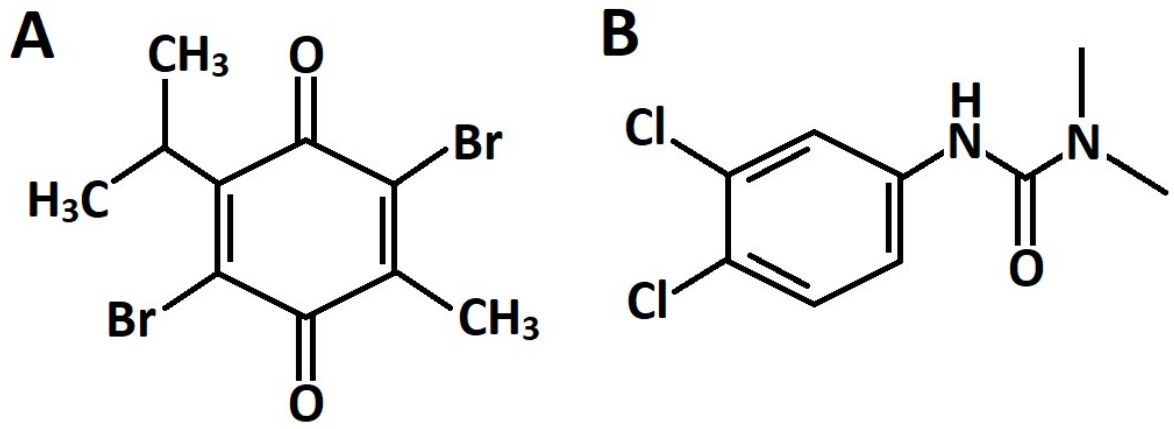


Figure 4.2: Chemical structures of DBMIB (A) and DCMU (B).

4.1.2.3. The herbicide DCMU targets PSII

3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) is also an analogue of plastoquinone (figure 4.3B). However, unlike DBMIB, DCMU specifically blocks the Q_A binding site at PSII (Moreland, 1980). DCMU has been extensively used for analyses that require inhibiting PSII or an oxidised plastoquinone pool, and due to its exclusivity for PSII, has been used to eliminate linear electron flow in favour of CEF, which does not require PSII (Slovacek *et al.*, 1978). The observations that DCMU inhibits linear electron flow only and that DBMIB inhibits both linear and cyclic electron flow have implications for the state transition (qT) stage of NPQ (Finazzi *et al.*, 1999). It was found that the formation of state 1 (more LHC around PSII rather than PSI) was inhibited by DCMU and DBMIB, yet state 2 formation (more LHC around PSI) was only inhibited by DBMIB. Therefore any difference in response to the two herbicides could be related to state transitions, or the interaction between cyclic and linear electron flow.

4.1.3. *Chlamydomonas reinhardtii* can grow photoautotrophically and heterotrophically

One feature that makes *C. reinhardtii* a useful model organism for photosynthesis is the ability to grow photoautotrophically, heterotrophically or mixotrophically (Sager and Granick, 1953). *C. reinhardtii* cultures grown under mixotrophic conditions with acetate have been observed to grow faster, with higher chlorophyll and protein contents than those grown photoautotrophically (Lalibertè and de la Noüie, 1993). The acetate is imported into the cell, converted to acetyl-coenzyme A, where it can be metabolised to produce reducing equivalents (Johnson and Alric, 2012). Therefore the acetate can complement the energy harvested during photosynthesis, and even drive ATP synthesis through transfer of electrons from NADPH to plastoquinone (Mus *et al.*, 2005, figure 4.1). This would suggest that *C. reinhardtii* cultures grown mixotrophically need to rely less on light as a source of energy.

Interestingly, mixotrophic growth of *C. reinhardtii* has been observed to have an effect on PSII kinetics and consequently ROS production. The contribution to the reduced plastoquinone pool of acetate metabolism without the need for PSII results in a transient LHC transition from PSII to PSI as the cells enter state 2 (Endo and Asada, 1996). Acetate can also compete with the essential water oxidation cofactor chloride at the water-evolving complex of PSII (Kühne *et al.*, 1999), thus decreasing the rate of reduction of the acceptor tyrosine. A thermoluminescence study has also found that acetate can bind to the Q_A^- site of PSII, which allows for more stable reduction of plastoquinone and a decrease in singlet oxygen production (Roach *et al.*, 2013). Therefore the

photoprotective mechanisms of *C. reinhardtii* can be affected by the presence or absence of acetate from the culture medium.

4.1.4. Knocking out Cytochrome c_{6A} could affect *C. reinhardtii* growth under stress conditions

The recent generation of *C. reinhardtii* CRISPR mutant lines with the cytochrome c_{6A} gene knocked out allows for potential growth phenotypes to be observed under different stress conditions (A. Ferenczi *et al.*, 2017). Three putative knockout lines were kindly provided by Attila Molnar and Aron Ferenczi. It was first necessary to confirm the nature of the mutations in the three strains. By observing how cytochrome c_{6A} knockout strains grow under fluctuating light conditions and growing photoautotrophically after a period of darkness, one can observe whether cytochrome c_{6A} is involved in the rapid response to changing light intensities. By observing growth under constant light that is either very low or high in intensity, one can observe the role cytochrome c_{6A} might play in longer term adaptation to certain light conditions.

Growth of the cytochrome c_{6A} knockout mutant lines in the presence of stress-inducing chemicals may also provide insight into which aspects of photosynthesis cytochrome c_{6A} is involved in. Exposure to RB would give insight into cytochrome c_{6A} interaction with singlet oxygen stress, and exposure to DCMU and DBMIB would give insight into cytochrome c_{6A} interaction with PSII and the plastoquinone pool. Also, a difference in response to DCMU and DBMIB could provide information on cytochrome c_{6A} interaction with state transitions and cyclic and linear electron flow. Performing the growth curves in both TP and TAP media will also allow observation of any interaction between mixotrophic growth and cytochrome c_{6A} function.

4.1.5 Chlorophyll fluorescence can indicate the efficiency of photosynthesis and photoprotection

Another method to measure the photoprotective response of *C. reinhardtii* would be through chlorophyll fluorescence analysis of quenching. When a chlorophyll molecule in the light harvesting complexes is excited by absorption of a photon the energy can have one of three fates. The energy can be passed on to another molecule through resonance energy transfer (usually to drive photosynthesis), or the energy can be dissipated as either heat or fluorescence (Krause, 1991; Maxwell and Johnson, 2000). As these three processes act in competition, one can measure the chlorophyll fluorescence of a photosynthetic organism to gain information about the efficiency of photosynthesis and heat dissipation in photoprotection (Kautsky *et al.*, 1960).

When a photosynthetic organism is kept in the dark, the PSII reaction centres eventually enter an 'open' state. This means that the Q_A downstream of the P680 reaction centre is able to receive an electron and electron transfer through that PSII complex, and consequently energy transfer to that PSII, can occur. When the organism is shifted to the light, the Q_A in the PSII complex is reduced and the reaction centre becomes 'closed'. A closed PSII reaction centre cannot receive more energy from peripheral chlorophyll until Q_A is oxidised by plastoquinone, therefore the dissipation of energy (fluorescence yield) increases. This all occurs in the timescale of around 1 second. After continued exposure to light, the efficiency of electron transfer through the electron transport chain increases, thus increasing the proportion of 'open' reaction centres at PSII and decreasing the amount of fluorescence. This is known as photochemical quenching. There is also an increase in the amount of photoprotective measures whereby heat dissipation of excited chlorophyll molecules is increased, known as non-photochemical quenching (NPQ). Both photochemical quenching and non-photochemical quenching will decrease the fluorescence yield of chlorophyll, and therefore using a specific light scheme (figure 4.17) the contribution of photochemical and non-photochemical quenching to photosynthetic efficiency can be measured.

4.1.5.1 Non-photochemical quenching has short and long term components that can be analysed through fluorescence

Non-photochemical quenching of high energy chlorophyll is a collection of photoprotective processes that prevent accessory chlorophyll from being in the excited state too long in high light conditions, as this leads to ROS production and stress (Krieger-Liszkay, 2005). In *C. reinhardtii*, non-photochemical quenching has several components, covering a response time from a few seconds to hours and days (Erickson *et al.*, 2015), which is reviewed in the introduction to this thesis.

Both photochemical and non-photochemical quenching can be analysed through chlorophyll fluorescence studies (Maxwell and Johnson, 2000). Three parameters of photochemical quenching: quantum yield of PSII photochemistry (ΦPSII); proportion of open PSII (qP); and the maximum quantum yield of PSII (F_v/F_m), give an insight into the overall efficiency of PSII-driven photosynthesis. ΦPSII gives an estimate of the proportion of energy from photon absorbance used in PSII for charge separation and photochemistry, and can be used as an indicator for linear electron transport rate. The parameter qP gives an indication as to how many of the PSII are available and in the 'open' state during a period of illumination, using the proportion of fluorescence levels given after dark adaptation and a period of actinic light. The third photochemical quenching parameter, F_v/F_m, links

qP and ΦPSII, and gives an indication of the theoretical efficiency of PSII if they were all in the 'open' state. In all of these parameters, a higher value indicates more efficient photochemical quenching, and therefore a more efficient rate of linear electron transport. Furthermore, chlorophyll fluorescence measurements made after a period of strong illumination can allow for an estimation of the contributions of fast-response and slow-response NPQ to photoprotection (Walters and Horton, 1991; Maxwell and Johnson, 2000).

Analysing the interaction between cytochrome c_{6A} and chlorophyll fluorescence by measuring how the CRISPR-cpf1 cytochrome c_{6A} knockout lines undergo photochemical and non-photochemical quenching will provide valuable information about the possible role of the former in photoprotection.

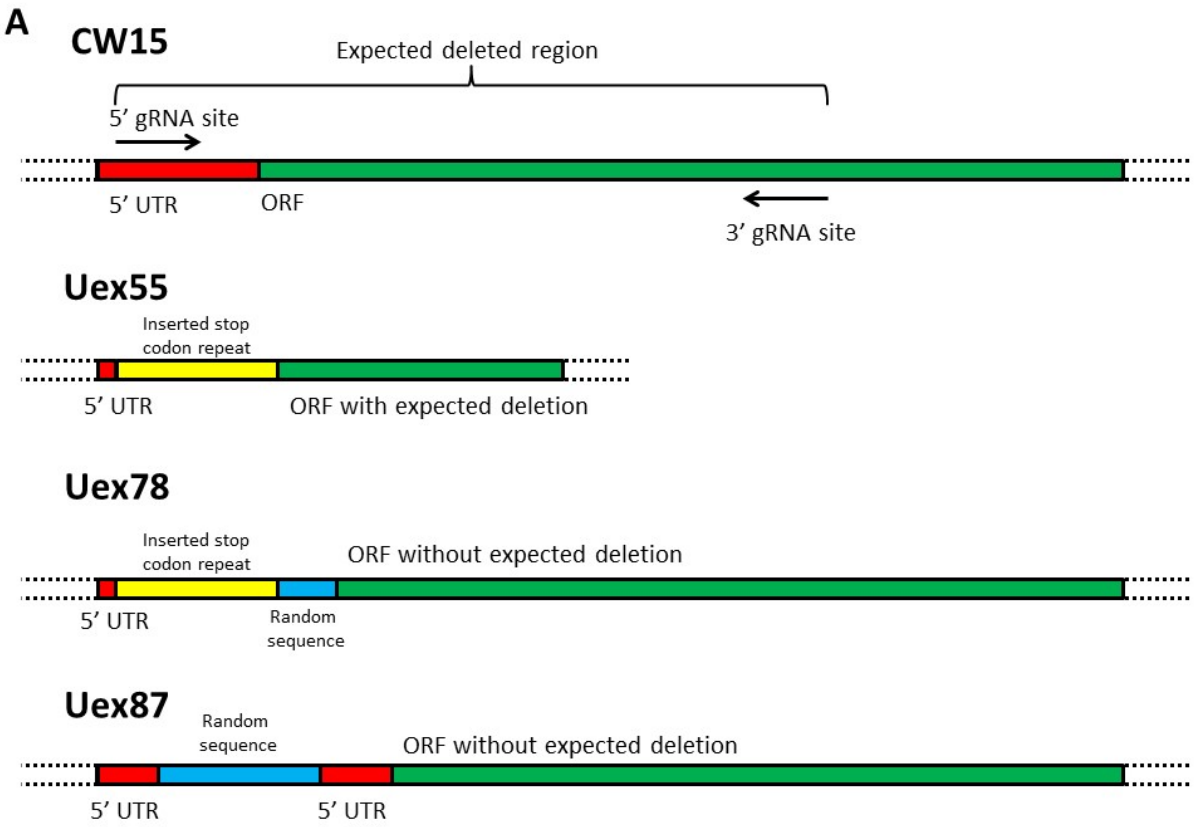
4.1.6 Aims of the study

- Analyse the mutations generated by CRISPR-Cpf1
- Determine if knocking out cytochrome c_{6A} in *C. reinhardtii* affects growth under extreme light conditions.
- Determine if knocking out cytochrome c_{6A} in *C. reinhardtii* affects growth when exposed to exogenous singlet oxygen.
- Determine if knocking out cytochrome c_{6A} in *C. reinhardtii* affects growth when exposed to PSII herbicides.
- Explore any interaction between cytochrome c_{6A} and heterotrophic growth.
- Analyse the efficiency of photosynthetic electron transfer in cytochrome c_{6A} mutant lines.
- Explore the NPQ properties of cytochrome c_{6A} mutant lines.

4.2. Results

4.2.1. Confirmation of CRISPR cytochrome c_{6A} knockout strains

Putative cytochrome c_{6A} knockout mutant lines Uex55, Uex78 and Uex87 were created using a CRISPR-Cpf1 system by Aron Ferenczi at the Molnar laboratory, University of Edinburgh. The background strain was the cell wall-deficient line cw15 (Davies and Plaskitt, 1971). The system used guide RNAs (gRNAs) that targeted a region between the 5' UTR and exon 3 of the cytochrome c_{6A} gene (*CYC4*). In conjunction with the repair template, the CRISPR-Cpf1 complex would be expected to delete this region and insert a short sequence containing stop codons in all reading frames in its place to disrupt translation of any truncated mRNA produced (figure 4.4). Whole genome sequencing (WGS) was performed on the mutant lines to confirm this genetic editing. Uex55 showed the expected deletion and stop codon repeat insertion, with no off-target modifications detected. However, Uex78 and Uex87 demonstrated more complicated sequences than expected. Uex78 was found to have the expected stop codon insertion within the 5'UTR, which was followed by a short region of the repair template and then an unrecognised DNA sequence, then the full coding region of *CYC4* with no deletions. The novel sequence totalled 39 nucleotides and replaced 36 nucleotides of the 5' UTR. It is not clear if the the 5'UTR being deleted would have an effect on transcription of *CYC4*, or transcript stability. The deleted 5' region ended three nucleotides before the start codon and immediately before the putative Kozak sequence of AAC (Cross, 2016). It is not clear if this would affect translation, but any mRNA translated would yield intact cytochrome c_{6A}. Uex87 was found to have an insertion of random DNA sequence in the 5'UTR , followed by the full coding region of *CYC4* with no deletions. In this mutant the 5'UTR might be disrupted by the insertion and a deletion of 20 nucleotides, and so if transcribed could still form a fully functional mRNA and subsequent protein. It appeared that the Kozak consensus sequence would not be disrupted by the insertion.


A
CW15
Expected deleted region
5’ gRNA site
5’ UTR
ORF
3’ gRNA site
Uex55
Inserted stop codon repeat
5’ UTR
ORF with expected deletion
Uex78
Inserted stop codon repeat
ORF without expected deletion
5’ UTR
Random sequence
Uex87
Random sequence
ORF without expected deletion
5’ UTR
5’ UTR

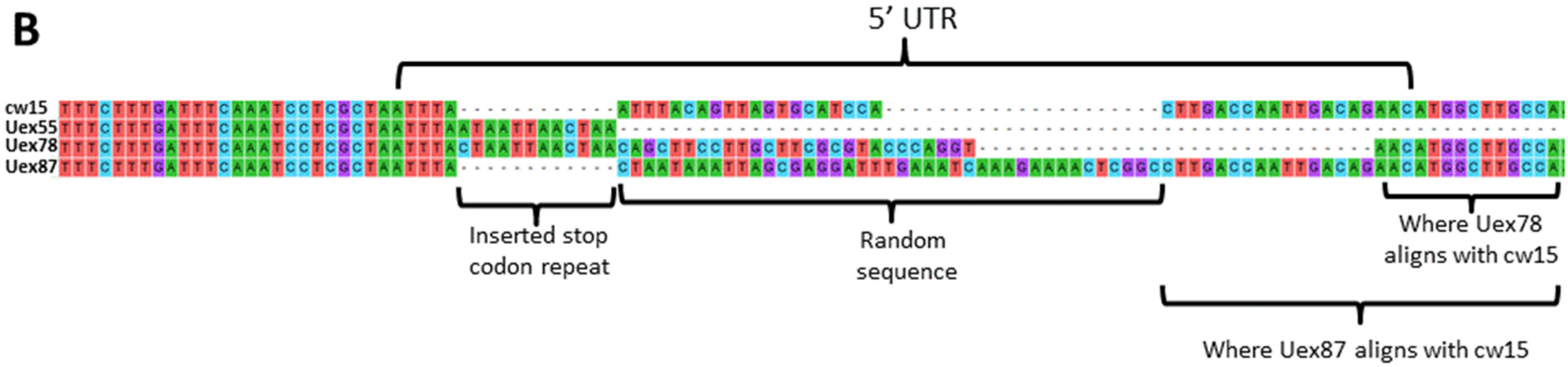

B
5’ UTR
cw15
Uex55
Uex78
Uex87
Inserted stop codon repeat
Random sequence
Where Uex78 aligns with cw15
Where Uex87 aligns with cw15

Figure 4.3: Comparison of the CYC4 region of the CRISPR-cpf1 knockout mutant lines and the background line (cw15) as determined through whole genome sequencing. A) Overview of the CYC4 region in the 4 strains. The wild-type form of CYC4 (cw15) showed a 5'UTR (red) followed by the open reading frame (ORF, green). In Uex55 most of the 5'UTR and ORF up to exon 3 was deleted, and a stop codon repeat sequence (yellow) had been inserted. In Uex78 most of the 5' UTR had been deleted and replaced with the inserted stop codon sequence and an unrecognised random sequence (blue). The ORF in Uex78 was unedited. In Uex87 the 5'UTR had a smaller deletion than in Uex55 and Uex78 but also contained an insertion of a random sequence, and the ORF was unedited. B) Sequence alignment of the CYC4 5'UTR and start of the ORF in the four strains. The four bases are coloured in each sequence, and specific features of the sequences are highlighted.

To determine whether transcript levels of cytochrome c_{6A} were affected by these modifications, RT-PCR was performed on RNA extracted from each mutant line and cw15 (figure 4.3). No transcripts were detected for cytochrome c_{6A} in any of the three mutant lines, therefore suggesting that transcription or transcript stability of *CYC4* had been affected. There is a possibility that cytochrome c_{6A} was still being transcribed in Uex78 and Uex87 at a level lower than detection, as the wild-type band was faint. These data show that Uex55 appears to be a true knockout mutant of cytochrome c_{6A}, with Uex78 and Uex87 being potential knockdown mutants.

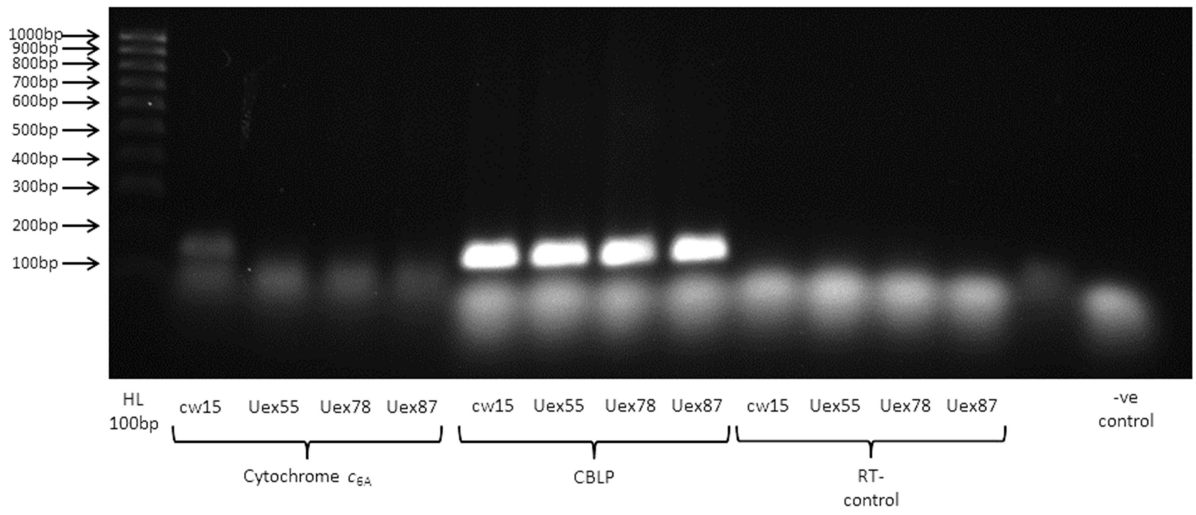


Figure 4.4: RT-PCR showing transcripts of cytochrome c_{6A} in the three CRISPR-Cpf1 mutant lines and cw15. CBLP PCR targets a constitutively expressed gene and was used as a positive control. The expected product size of cytochrome c_{6A} and CBLP transcript were around 180 and 150 base pairs respectively. HL stands for hyperladder. The RT- control used CBLP primers on extracted RNA that had not been reverse transcribed. The negative control (far right) contained the PCR reaction mixture with no cDNA.

Therefore all the mutant lines have been presented in this study, but it should be noted that Uex55 is the only definitive knock out line, with Uex78 and Uex87 being potential knockdowns of cytochrome c_{6A} transcript levels and translation. Given the low levels of cytochrome c_{6A} expression under normal conditions, it did not seem helpful to try to use western analysis to test for protein levels.

4.2.2. Growth curves under differing light intensities

4.2.2.1. Initial plate-spot tests of differing light conditions and Rose Bengal

A preliminary spot-plate test of the effects of light intensity and RB on cytochrome c_{6A} knock out or knockdown lines was performed. In the experiment, 20 µL of cultures of each cytochrome c_{6A} Uex line and cw15 at an OD_{750} of 0.5 were placed on solid TP or TAP agarose plates and the plates placed under different conditions for 13 days (figure 4.6). Plates were either placed under standard light (SL, 40 µE m^{-2} s^{-1}), high light (HL, 400 µE m^{-2} s^{-1}) or complete darkness. Half of the plates also contained 10 µM RB to allow comparison of how light intensity may interact with singlet oxygen stress.

In complete darkness, the TP plates showed no colonies as the minimal media only allowed photoautotrophic growth. TAP plates in total darkness showed colonies with no noticeable difference between 0 µM RB and RB. As RB only generates singlet oxygen in light, it is likely that a negligible amount of singlet oxygen was generated in the complete darkness plates, further supported by the fact that those plates had retained their pink colour when under SL and HL the RB in the plates have lost their pink colour. Under SL there appeared to be slower growth of the knockout or knockdown lines compared to cw15 on the TAP RB plate, especially compared to the TAP 0 µM RB plate. This could indicate that a lack of cytochrome c_{6A} was disadvantageous under singlet oxygen stress during mixotrophic growth. There was potentially a slower growth of the knockdown lines Uex78 and Uex87 compared to cw15 on the TP RB plate, but this observation was difficult to determine with certainty from this qualitative experiment.

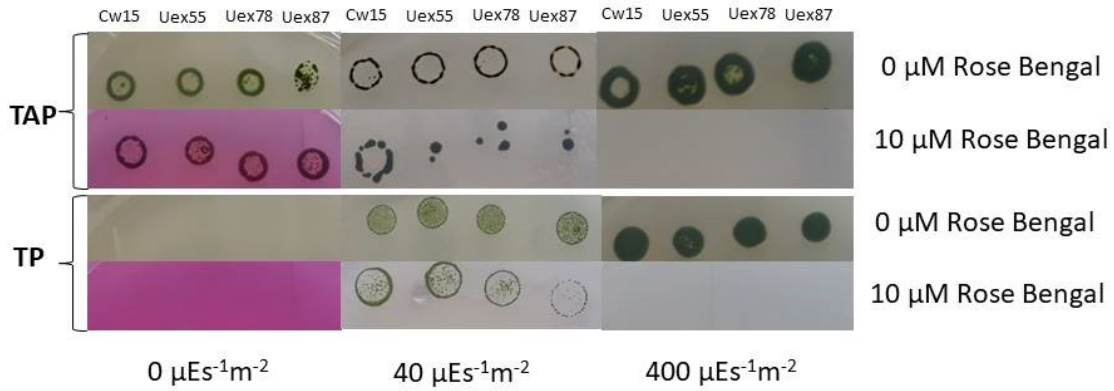


Figure 4.5: Spot tests of Uex55, Uex78 and Uex87 and the background strain cw15 under varying light intensities on agar plates containing varying concentrations of Rose Bengal. This was performed with both TP agar plates and TAP agar plates. The photographs were taken after 13 days under the respective conditions.

Under HL no growth on either TAP or TP plates had occurred in the presence of RB. This is likely to be because the cultures could not survive under both high light and extra singlet oxygen stress. There appeared to be no noticeable difference between the Uex strains and cw15 under HL stress on either TAP or TP media in the absence of RB.

4.2.2.2. Growth curves under constant low or high light conditions

To observe the effect of light intensity on the cytochrome c_{6A} knockout and knockdown lines quantitatively, liquid cultures of each line and cw15 were grown under SL, HL and low light (LL, 10 µE m^{-2} s^{-1}; figure 4.7). Under SL, there appeared to be no notable difference in growth between the Uex lines and cw15, supported by the calculated growth rates shown in figure 4.8. There was a slightly faster growth of Uex78 observed under SL in TAP medium; however, the other two mutant lines do not show such a phenotype.

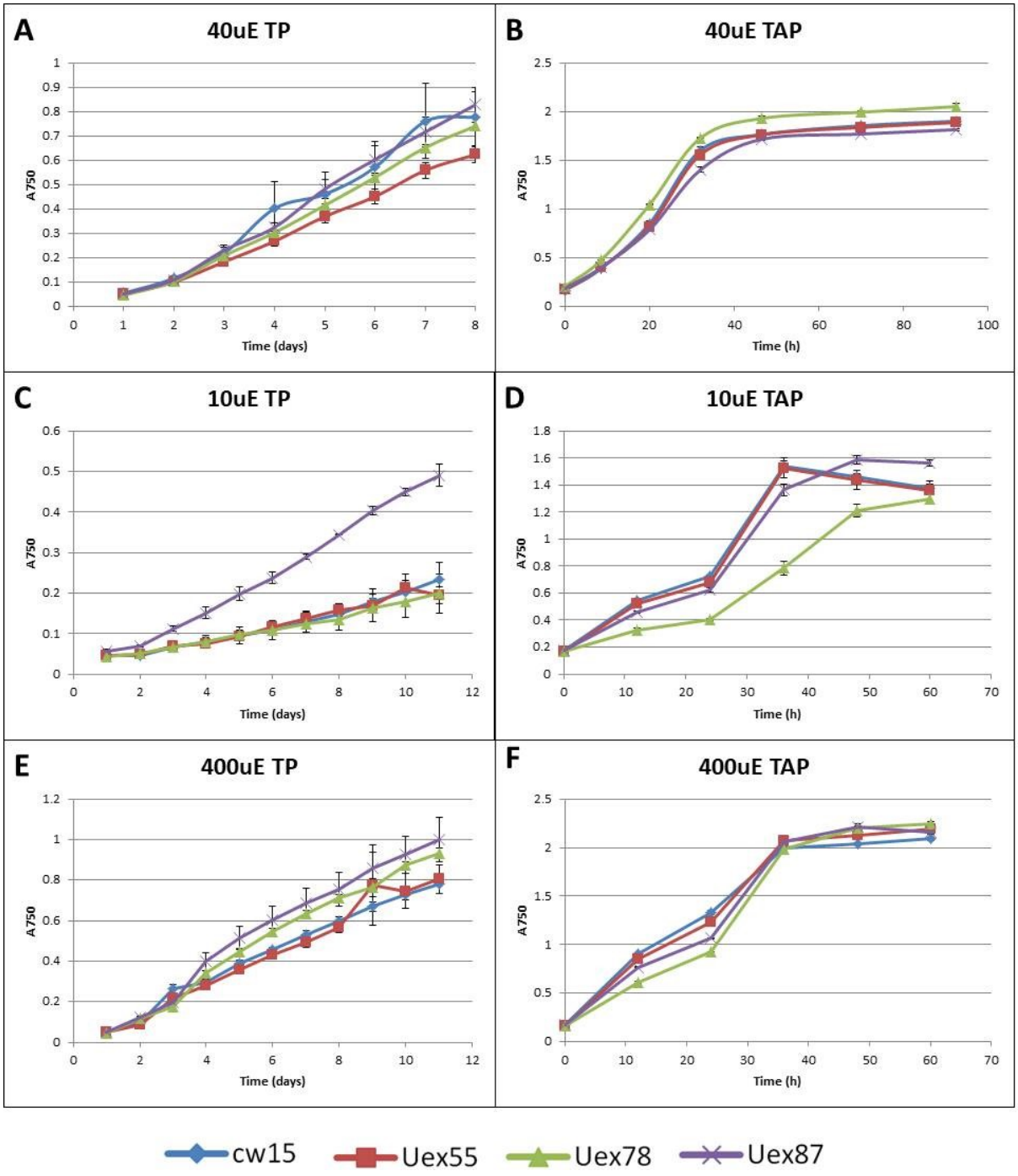


Figure 4.6: Growth curves for cytochrome c_{6A} mutant lines and cw15 under standard light (40 µE m^{-2} s^{-1}), low light (10 µE m^{-2} s^{-1}) and high light (400 µE m^{-2} s^{-1}) in both TP and TAP media. Error bars show the standard error for three technical replicates.

Under low light (LL, 10 µE m^{-2} s^{-1}), in TP medium Uex87 had a significantly higher growth rate ($p<0.001$) than cw15, but this was not observed in either Uex55 or Uex78. Therefore it is uncertain if this growth phenotype was due to the decrease in cytochrome c_{6A} expression. Growth of Uex78 was

significantly lower than cw15 ($p<0.001$) in TAP media under LL, but again this phenotype was not observed in the other two mutant lines.

Under HL conditions, there was no notable difference in growth phenotypes between the cytochrome c_{6A} knockout or knockdown mutants and cw15 in TP media. There was a potential slow growth phenotype for the Uex mutants relative to cw15 in TAP media under HL, at least in the early part of culture growth, which was supported by significant differences in growth rate between Uex78 and Uex87 from cw15 ($p<0.001$ for both lines), but the growth rate of Uex55 was not significantly lower ($p=0.343$). Therefore there could be a link between cytochrome c_{6A} and high light stress under mixotrophic conditions.

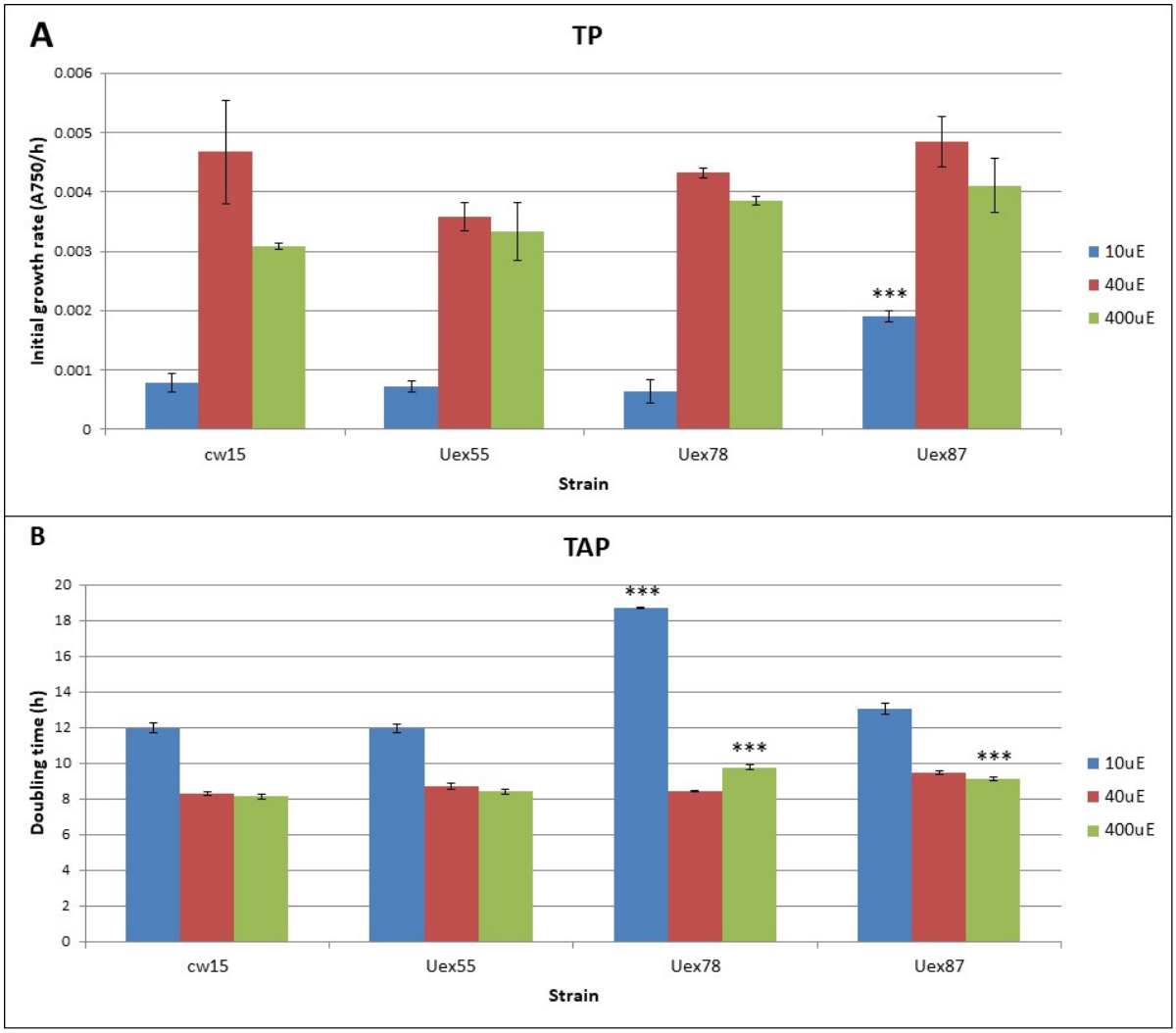


Figure 4.7: Growth rates (TP, A) and doubling time (TAP, B) for the cytochrome c_{6A} knockout mutants and cw15 under standard light, low light and high light intensities. Growth rates were calculated instead of doubling time for the TP curves because the cultures did not enter stationary phase and were thus difficult to estimate. The error bars represent standard error from three technical

*replicates. Asterisks denote significance in the Uex mutant lines relative to cw15 for the relevant condition as measured by ANOVA with the Tukey post-hoc test, where *, ** and *** represent a p-value < 0.05, < 0.01 and < 0.001 respectively.*

4.2.2.3. Growth curves under constant light following a period of darkness

To determine whether the transition from dark to light was affected by a lack of cytochrome c_{6A}, the Uex mutant lines and cw15 were exposed to constant 40 µE m^{-2} s^{-1} following a period of 48 h in total darkness (figure 4.9). There was a trend of diminished growth on illumination observed in all of the cytochrome c_{6A} mutant lines relative to cw15, although the differences were not statistically significant.

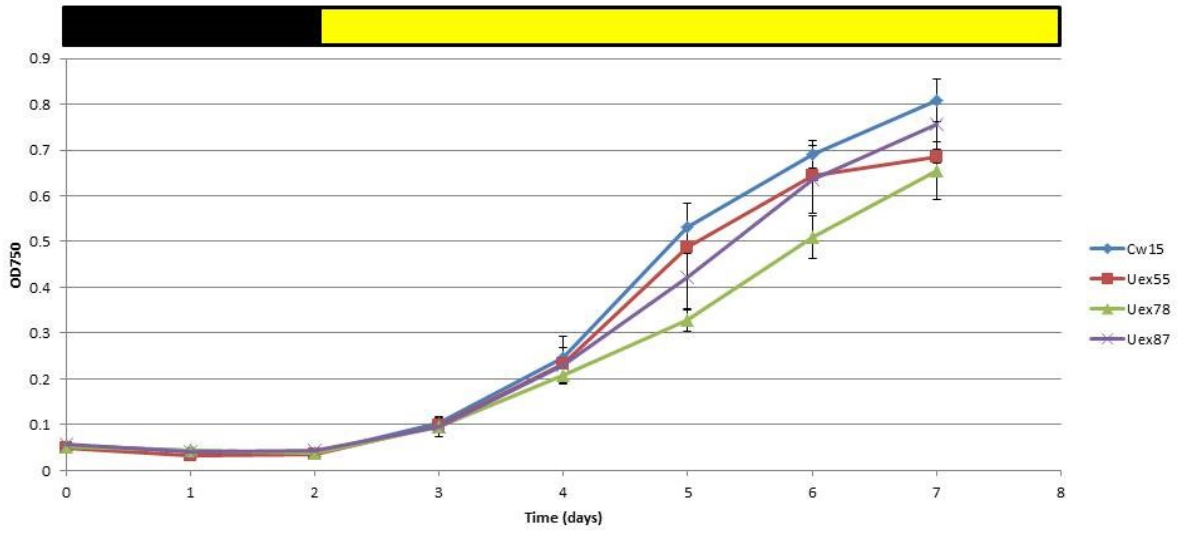


Figure 4.8: Growth curves for cytochrome c_{6A} mutant lines and cw15 in TP media under standard light after 48 h in darkness. Black and yellow bars at the top of the graph show the period of darkness and light respectively. Error bars show the standard error for three technical replicates.

The calculated growth rates from the period straight after the transition from dark to light also showed a trend of diminished growth in the Uex mutant lines (figure 4.10). However, only the growth rate of Uex78 was found to be statistically significantly lower than cw15 (p=0.029), with Uex55 and Uex87 not showing statistical significance (p=0.307 and 0.580 respectively). Therefore again there was a trend in diminished growth that was consistent among mutant lines relative to cw15, but it was not statistically significant.

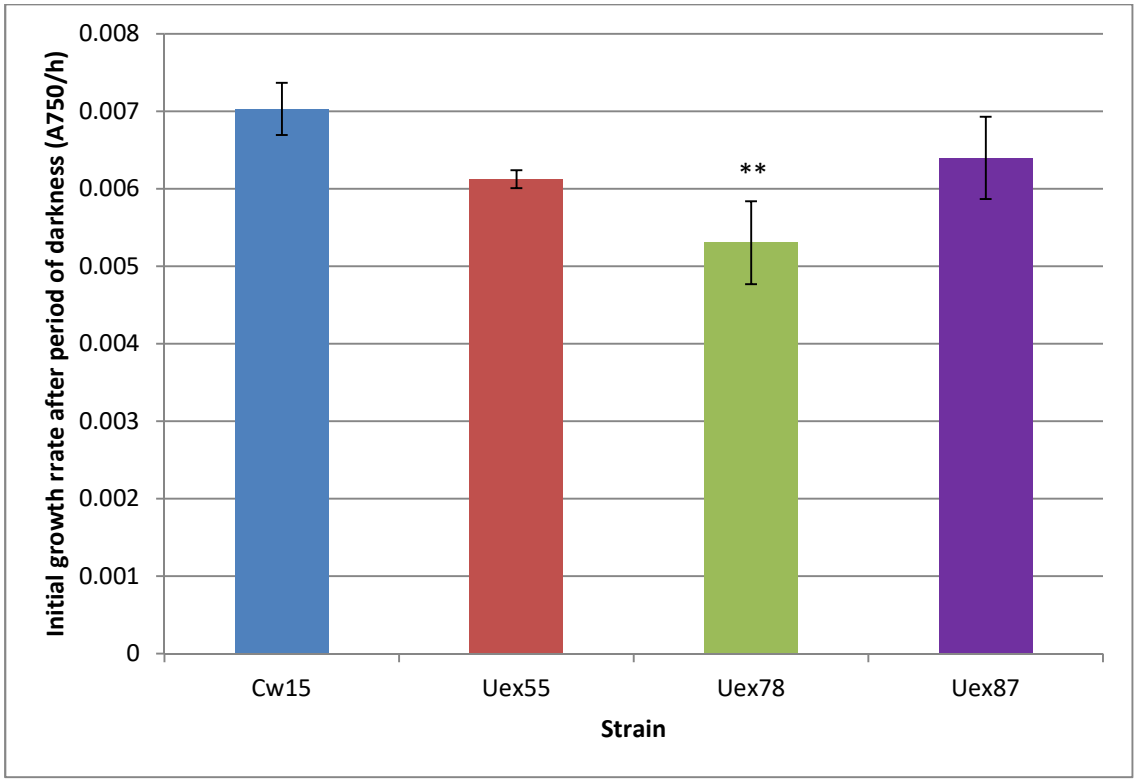


*Figure 4.9: Growth rates for the cytochrome c_{6A} mutants and cw15 under standard light following a period of 48 h in the dark. The error bars represent standard error from three technical replicates. Asterisks denote significance in the Uex mutant lines relative to cw15 for the relevant condition as measured by ANOVA with the Tukey post-hoc test, where *, ** and *** represent a p-value < 0.05, < 0.01 and < 0.001 respectively.*

4.2.2.4. Growth curves under fluctuating light

To investigate further the potential link observed between dark to light transitions and cytochrome c_{6A}, growth of cw15 and knockout mutant line Uex55 were measured during a diurnal cycle of 12 h at 40 µE m^{-2} s^{-1} followed by 12 h of total darkness in the Algem bioreactor (Algenuity; figure 4.11). As was observed for dark to light transition in figure 4.9 a trend was observed of slower growth in the knockout mutant relative to cw15 (figure 4.12). However, large differences in growth rates between biological replicates led to very large error bars and a statistically insignificant change in growth rate (p=0.181). Therefore again there is a possible link between cytochrome c_{6A} and the cycle between dark and light, but this is not certain.

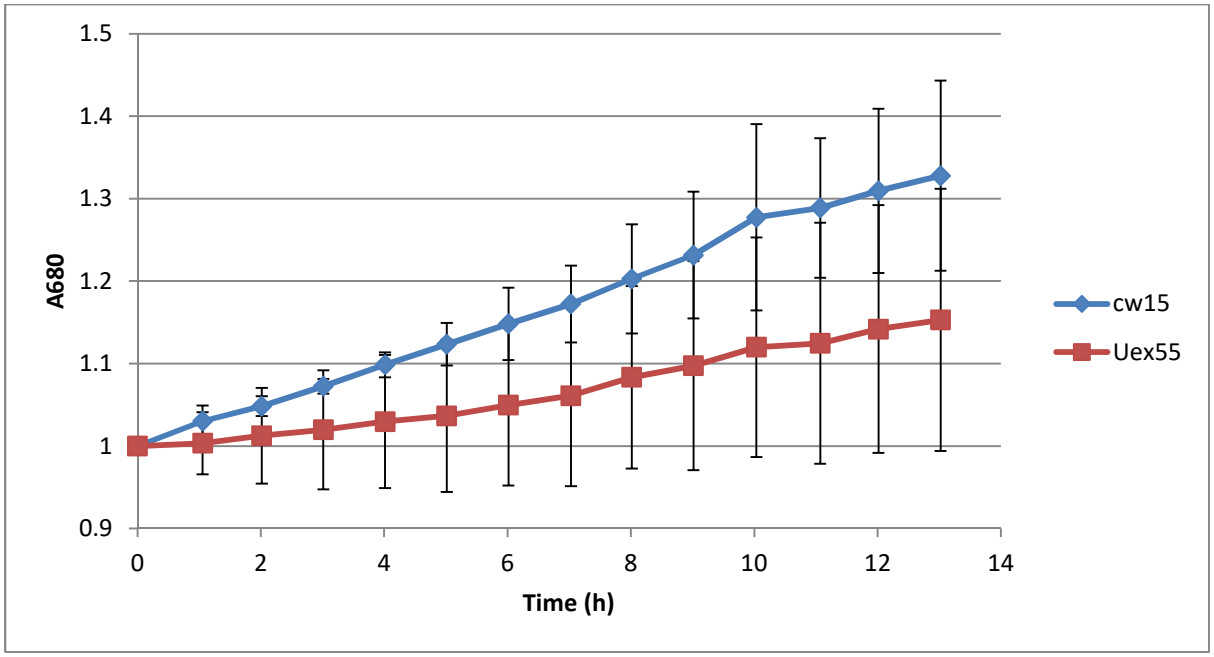


Figure 4.10: Growth curves of cytochrome c_{6A} knockout line Uex55 and cw15 under a diurnal cycle of 12 h standard light and 12 h darkness. Growth curves were normalised to point zero. Error bars represent standard error across three biological replicates.

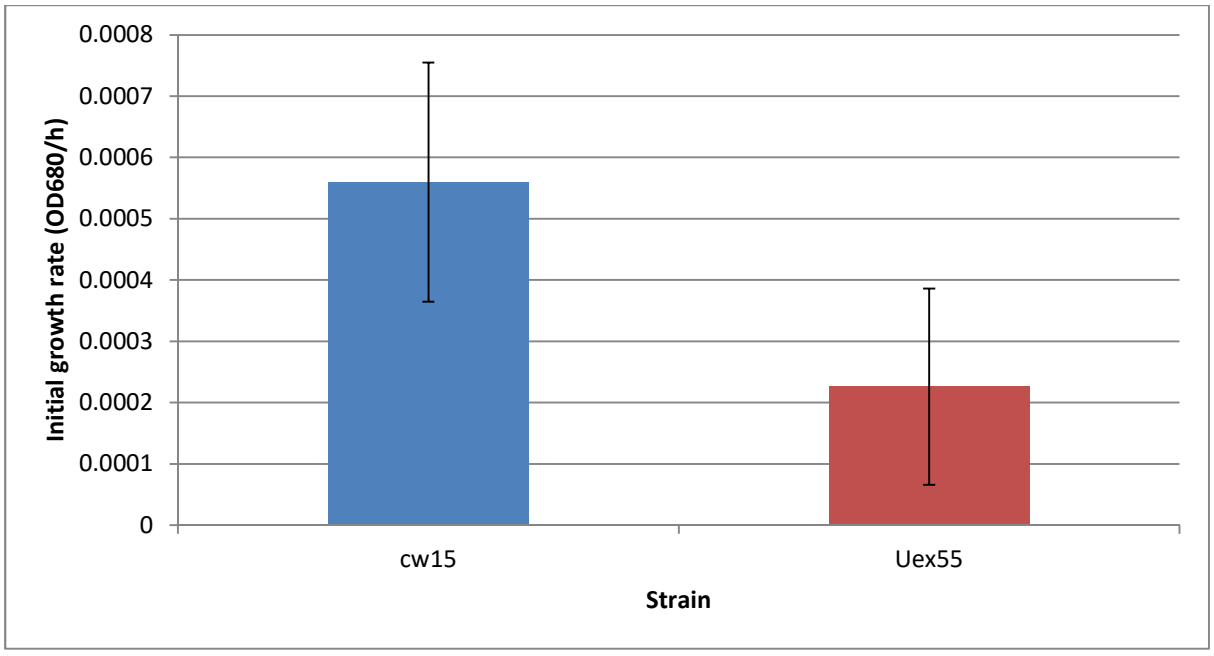


Figure 4.11: Initial growth rates for cw15 and Uex55 during exposure to the diurnal cycle of 12 h standard light and 12 h darkness. Error bars represent standard error across three biological replicates. Asterisks denote significance in the Uex mutant lines relative to cw15 for the relevant

*condition as measured by ANOVA with the Tukey post-hoc test, where *, ** and *** represent a p-value < 0.05, < 0.01 and < 0.001 respectively.*

4.2.3. Growth curves with exposure to stress-inducing chemicals

4.2.3.1. Growth curves upon Rose Bengal exposure

Following the preliminary experiment in section 4.2.1.1, growth of the cytochrome c_{6A} mutant lines was compared to cw15 in liquid cultures with 10 µM RB (figure 4.13). In TP medium, strain Uex78 demonstrated a slightly faster growth, supported by a higher calculated growth rate (figure 4.14). However, neither Uex55 nor Uex87 showed this faster growth. In TAP medium, all the Uex mutant lines appeared to have a longer lag phase before exponential growth relative to cw15. Interestingly, the calculated growth rates did not appear markedly different except for Uex87, which had a significantly longer doubling time than cw15 in TAP media under RB stress ($p = 0.044$). Therefore it seems that the time taken to adjust to the RB was longer in the cytochrome c_{6A} mutant lines than cw15, but the eventual growth rate once the adjustment had occurred was not different between the mutant lines and cw15. Therefore when grown mixotrophically it appeared that a lack of cytochrome c_{6A} led to a longer adaptation period to singlet oxygen exposure.

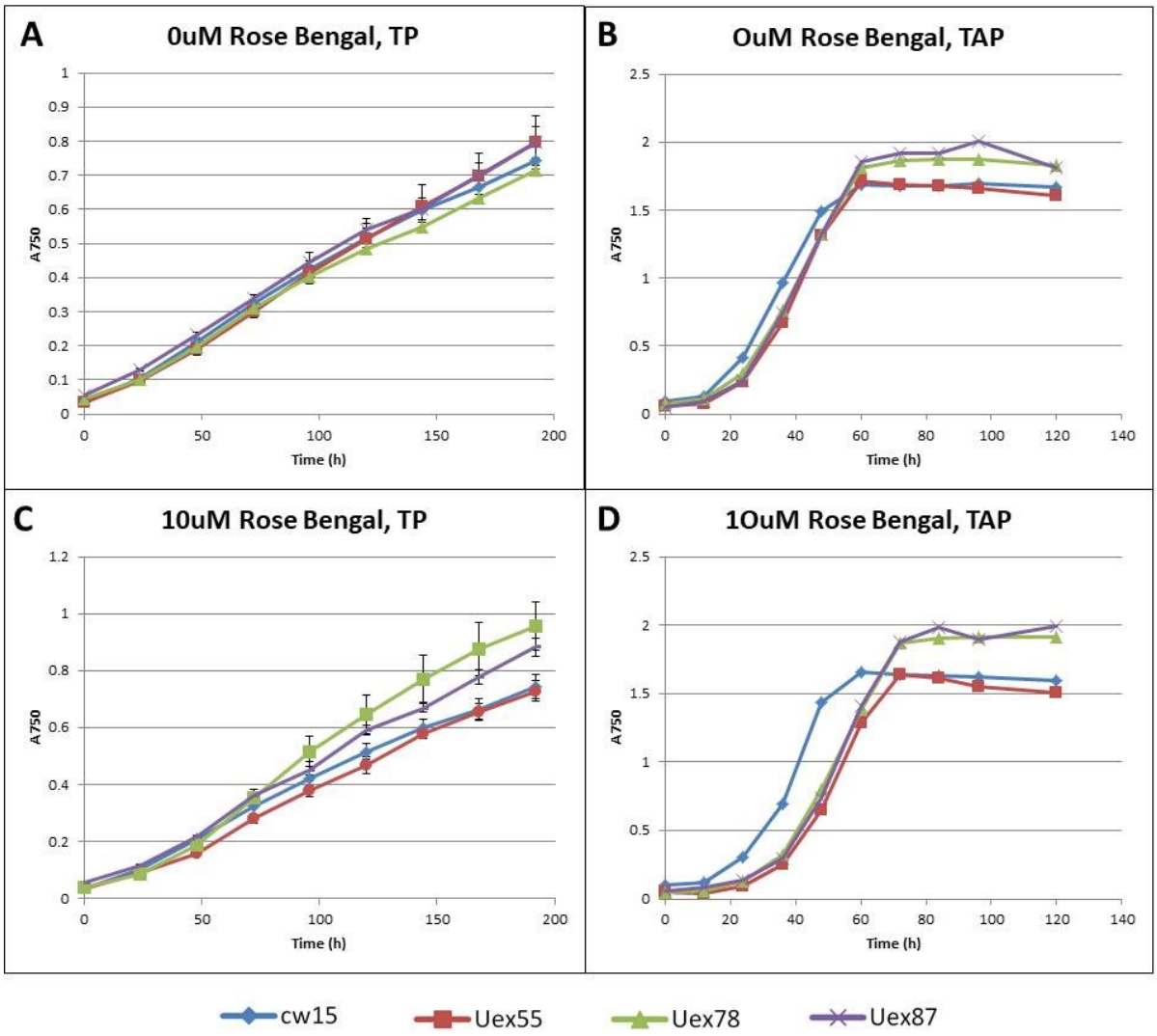


Figure 4.12: Growth curves for cytochrome c_{6A} mutant lines and cw15 upon exposure to 10 µM Rose Bengal in both TP and TAP media. Error bars show the standard error for three technical replicates.

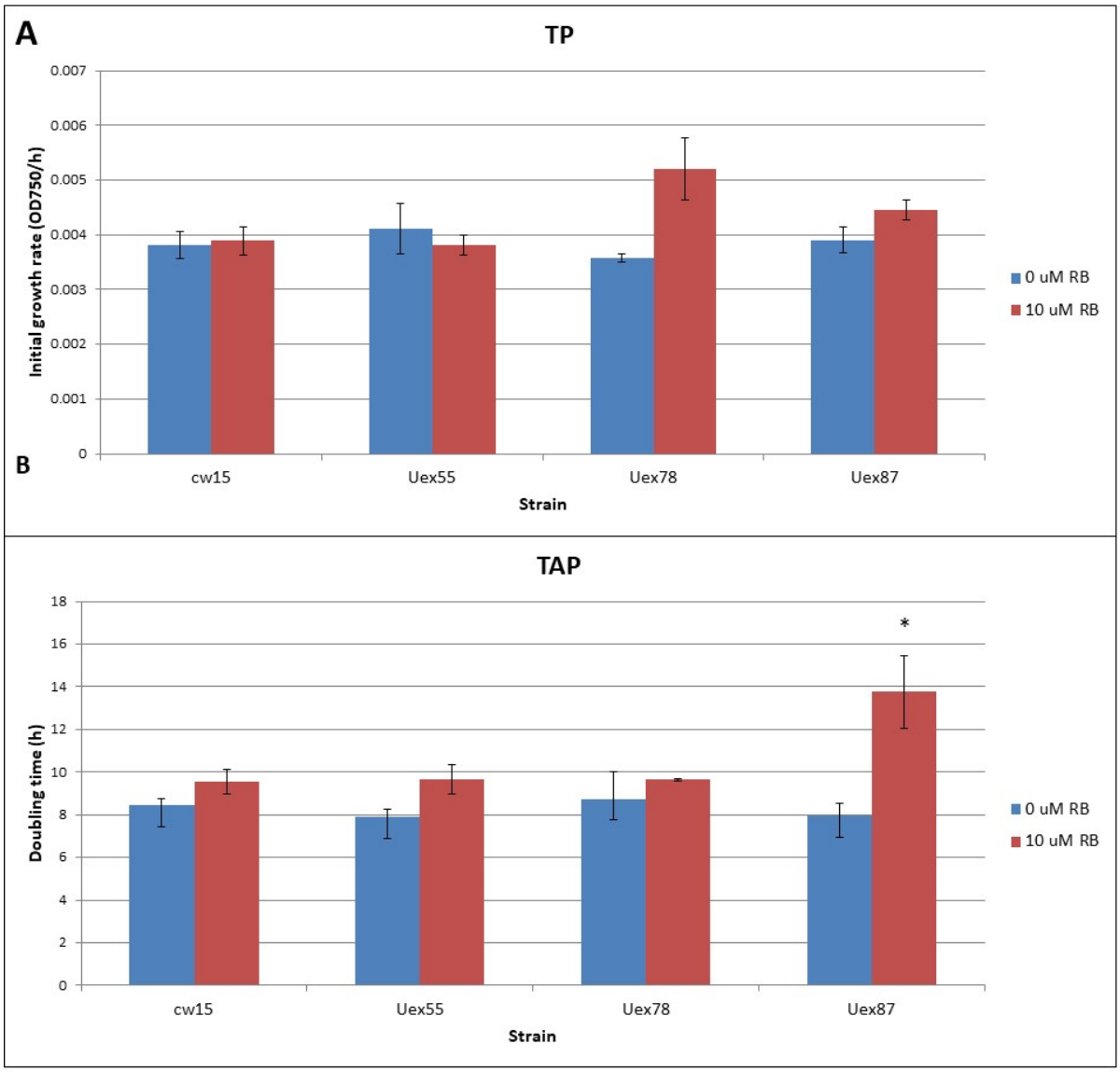


*Figure 4.13: Growth rates (TP, A) and doubling time (TAP, B) for the cytochrome c_{6A} knockout and knockdown mutants and cw15 upon exposure to 10 µM Rose Bengal. Growth rates were calculated instead of doubling time for the TP curves because the growth did not run to completion and were thus difficult to estimate. The error bars represent standard error from three technical replicates. Asterisks denote significance in the Uex mutant lines relative to cw15 for the relevant condition as measured by ANOVA with the Tukey post-hoc test, where *, ** and *** represent a p-value < 0.05, < 0.01 and < 0.001 respectively.*

4.2.3.2. Growth curves upon DBMIB and DCMU exposure

DBMIB and DCMU are potent inhibitors of photosynthesis, therefore preliminary small scale growth curves were determined at increasing concentrations of both herbicides in order to determine an

effective but non-lethal dose (figure 4.15). The data showed that 1 µM DBMIB and 0.1 µM DCMU would diminish growth of the strains without being lethal, and were therefore used in subsequent experiments.

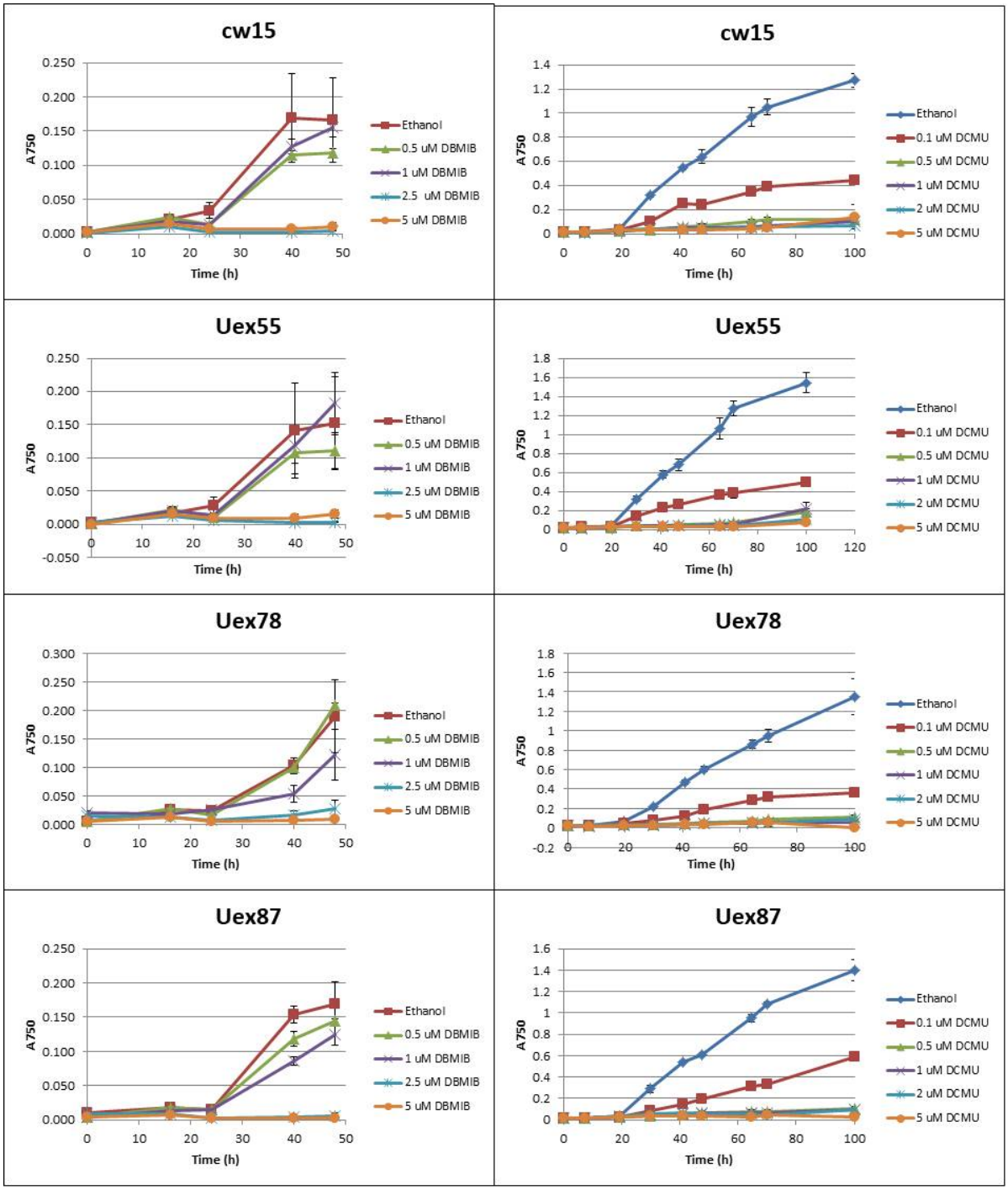


Figure 4.14: Preliminary growth curves of the cytochrome c_{6A} mutants and cw15 under different concentrations of DBMIB (left) and DCMU (right). Cultures were grown in 6 mL wells containing each herbicide dissolved in ethanol. The concentration of ethanol in each well was 0.01% v/v. Error bars represent standard error from three technical replicates.

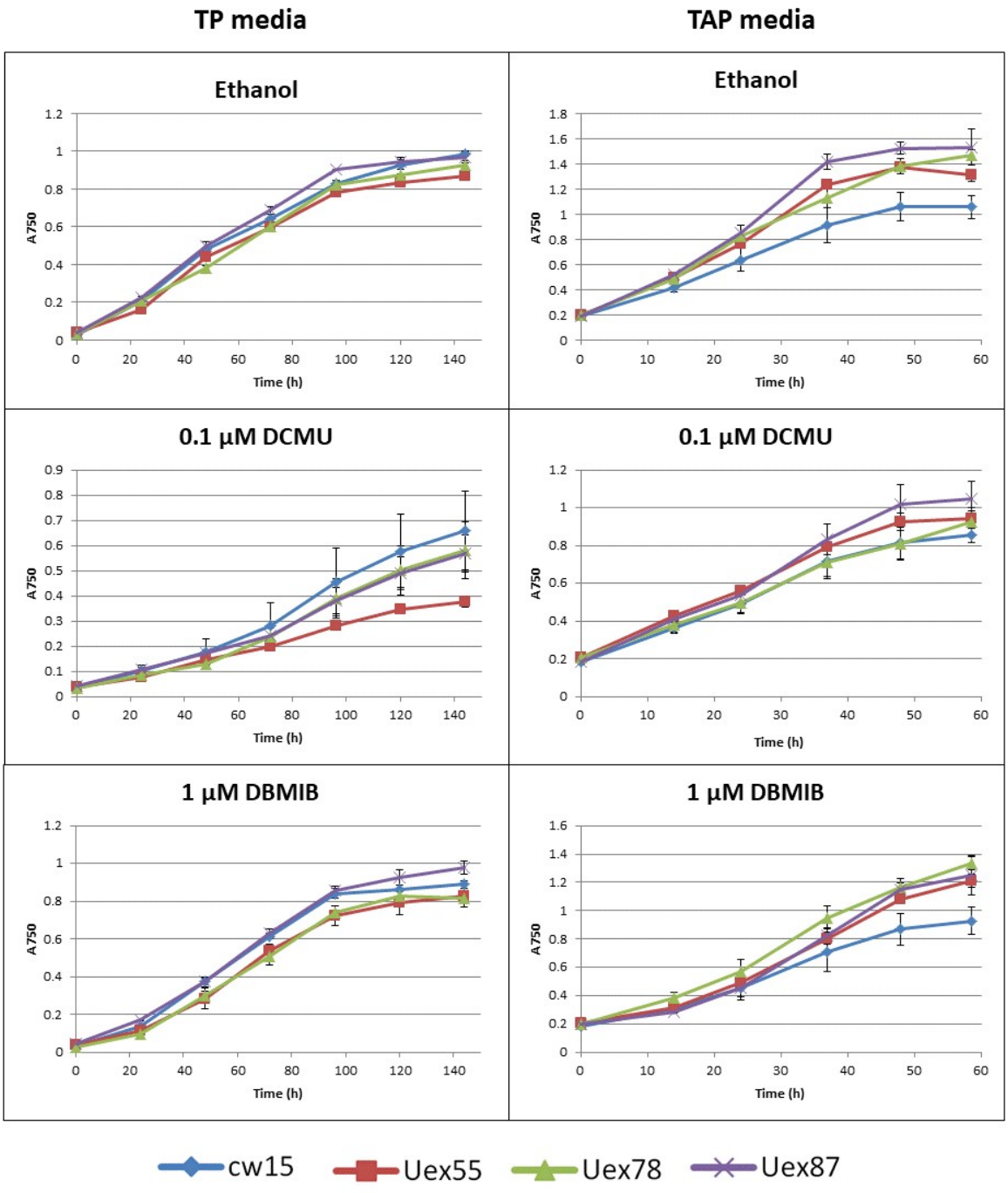


Figure 4.15: Growth curves for cytochrome c_{6A} mutant lines Uex55, Uex78 and Uex87, and cw15 upon exposure to either 0.1 µM DCMU, 1 µM DBMIB or equivalent volume of ethanol in both TP (left panels) and TAP media (right panels). Error bars show the standard error for three technical replicates.

The cytochrome c_{6A} mutant lines and cw15 were therefore exposed to non-lethal concentrations of DBMIB (1 μM) or DCMU (0.1 μM) in both TP and TAP media (figure 4.17). In TP media, the Uex mutants showed a potential trend for slower growth when exposed to either DBMIB or DCMU relative to cw15, which is supported by the trend seen in the calculated growth rates (figure 4.18). However, these trends were shown to be mostly statistically insignificant, with only the DBMIB TP difference between Uex55 and cw15 being significant at the 5% level (p=0.049). For the DCMU growth rates, the lack of significance seemed mostly to be due to the large error of the cw15 and Uex55 replicates. In TAP, there appeared to be no noticeable trend between the Uex mutant lines and cw15.

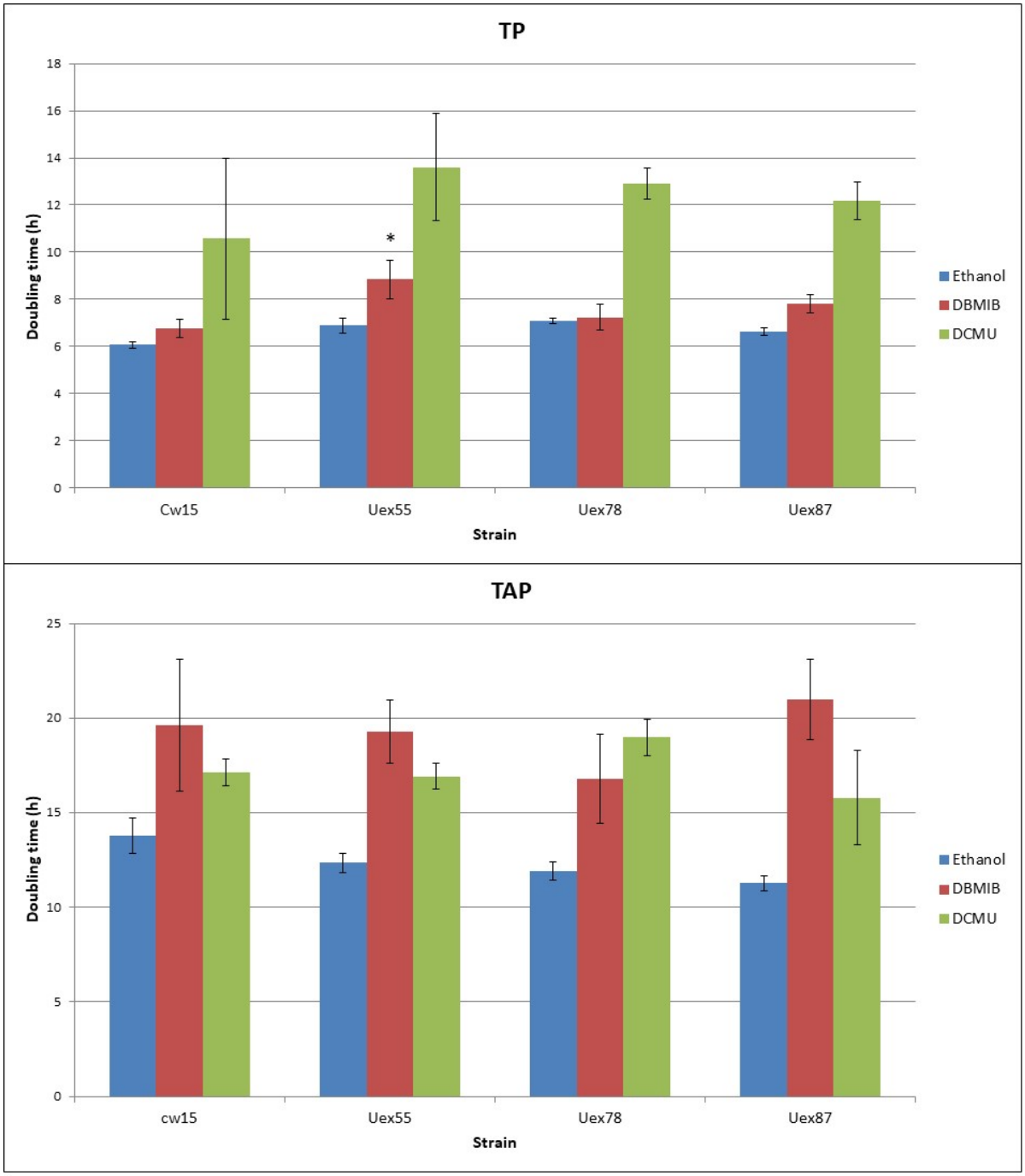


*Figure 4.16 Doubling times for the cytochrome c_{6A} mutants and cw15 upon exposure to either 0.1 µM DCMU, 1 µM DBMIB or equivalent volume of ethanol in TP and TAP media. The error bars represent standard error from three technical replicates. Asterisks denote significance in the Uex mutant lines relative to cw15 for the relevant condition as measured by ANOVA with the Tukey post-hoc test, where *, ** and *** represent a p-value < 0.05, < 0.01 and < 0.001 respectively.*

4.2.4. Summary of growth phenotypes of the cytochrome c_{6A} mutant lines

A summary of the growth phenotypes and their significance can be found in table 4.1.

Conditions	Figure	TP	TAP
Standard light (40 µE m^{-2} s^{-1})	4.7	No noticeable difference	No noticeable difference
High light (400 µE m^{-2} s^{-1})	4.7	No consensus among mutant lines	Trend of diminished growth for Uex mutants (only significant in Uex78 and Uex87, $p<0.001$)
Low light (10 µE m^{-2} s^{-1})	4.7	No consensus among mutant lines, Uex87 significantly faster initial growth ($p<0.001$)	No consensus among mutant lines, significantly slower growth in Uex78 ($p<0.001$)
48h in dark followed by 40 µE m^{-2} s^{-1}	4.9	Trend of diminished growth for Uex mutants, only Uex78 showed significant difference in growth rate ($p = 0.029$)	n/a
Fluctuating light (12 of 40 µE m^{-2} s^{-1} and 12 h of darkness)	4.11	Trend of diminished growth for Uex mutants (not significant)	n/a
10 µM Rose Bengal	4.13	No consensus among mutant lines	Trend of increased time spent in lag phase for Uex mutants, Uex87 had significantly diminished growth rate ($p = 0.044$)
0.1 µM DCMU	4.16	Trend of diminished growth for Uex	No noticeable difference

		mutants (not significant)	
1 µM DBMIB	4.16	Trend of diminished growth for Uex mutants , only Uex55 showed significant difference in growth (p = 0.049)	No noticeable difference

Table 4.1: Summary of the growth curve phenotypes observed in this chapter.

4.2.5 Use of chlorophyll fluorescence to measure photochemical and non-photochemical quenching

The three Uex mutant lines of *C. reinhardtii* were used in the chlorophyll fluorescent studies. In preliminary experiments, a basic chlorofluorescence trace was used under standard light conditions to gain an idea of photochemical and non-photochemical quenching with and without cytochrome c_{6A}. Subsequently, non-photochemical quenching in the Uex mutant lines was analysed and parsed further through relaxation analysis after exposure to high light.

4.2.5.1 A basic chlorofluorescence trace was used to estimate photochemical and non-photochemical quenching

The chlorophyll fluorescence properties of the cytochrome c_{6A} knockout mutant line Uex55, and the potential mutant knockdown lines Uex78 and Uex87, were determined using the CFImager from Technologica, and compared to that of the background strain cw15. This was achieved through measuring chlorophyll fluorescence for each line during a specific light scheme (figure 4.17). The specific quenching parameters described in section 5.1.6 can then be calculated using the fluorescence levels measured from the trace given for each mutant line (table 4.2).

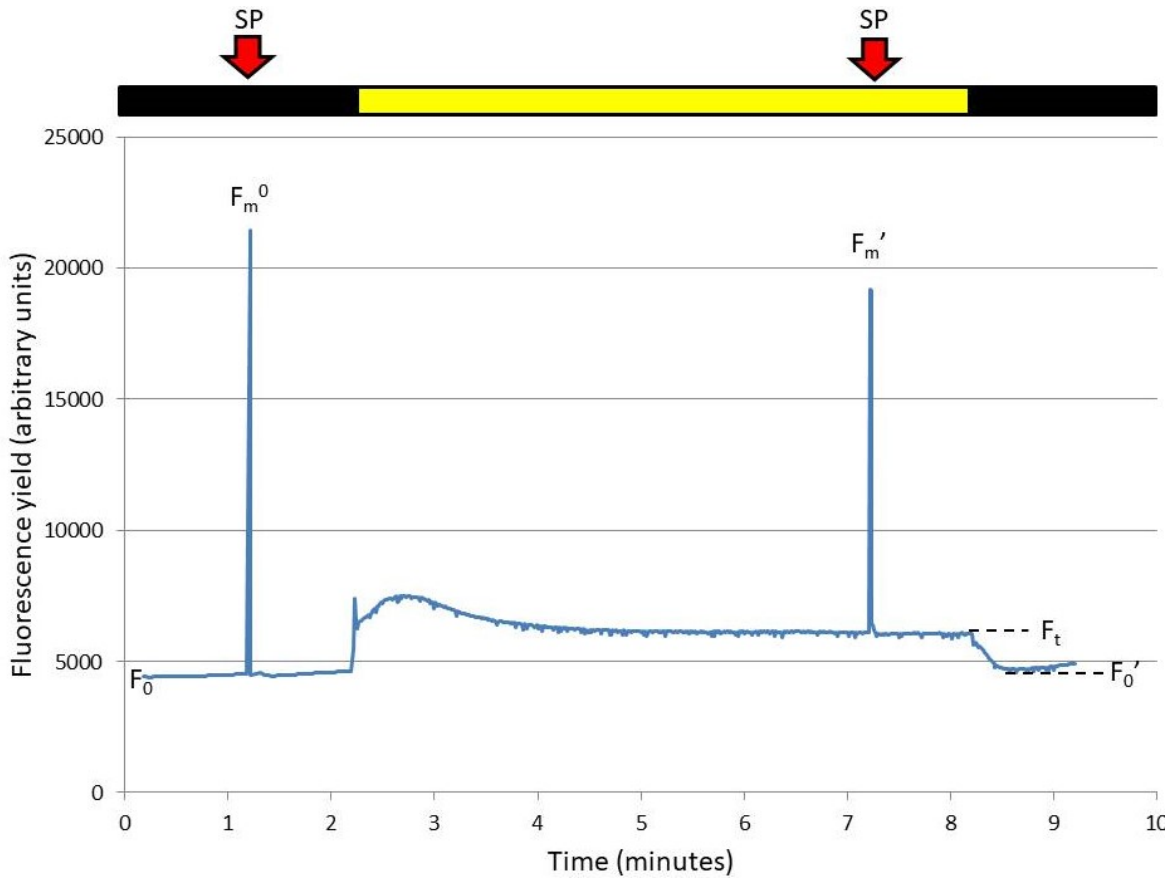


Figure 4.17: Raw chlorofluorescence trace from light regime performed to calculate the fluorescence parameters. After 10 min of dark adaptation, a measuring light (1.65 µE m^{-2} s^{-1}, black box) was turned on to measure the base level of fluorescence (F_0). After a minute in the dark a saturating pulse (SP) was applied to the sample, and the resulting fluorescence was measured (F_m^0). After a further minute an actinic light of 40 µE m^{-2} s^{-1} (yellow box) was turned on for 6 minutes. In the final minute of the actinic light period, a second SP was applied, and the resulting fluorescence was measured, giving a value of F_m'. The fluorescence level just before or after this pulse was taken as the F_t value. At the end of the 6 minutes the actinic light was switched off, and the lowest point of fluorescence of the trace in the following minute was taken as F_0'. All SPs were an intensity of 6172 µE m^{-2} $s^{-1}$for 800 ms, and the light used throughout was of 470 nm wavelength. Adapted from Maxwell and Johnson (2000).

The non-photochemical quenching value calculated was related to the effectiveness of photoprotective elements in the sample for dissipating energy as heat, as discussed in section 4.1.5. A higher value of NPQ indicates a more effective quenching ability of the sample.

Quenching parameter	Formula
Quantum yield (efficiency of charge separation) of PSII (ΦPSII)	$\frac{(F_m' - F_t)}{F_m'}$
Proportion of open PSII (qP)	$\frac{(F_m' - F_t)}{(F_m' - F_0')}$
Maximum quantum yield of PSII (F_v/F_m)	$\frac{(F_m - F_0)}{F_m}$ or $\frac{\Phi PSII}{qP}$
Non-photochemical quenching (NPQ)	$\frac{(F_m 0 - F_m')}{F_m'}$

Table 4.2: Formulae for calculating the quenching parameters from the fluorescence levels measured in the trace in figure 4.17. Taken from Maxwell and Johnson (2000).

The calculated parameters for the cytochrome c_{6A} mutants and cw15 are shown in figure 4.18. No significant difference was observed between the Uex lines and cw15 for ΦPSII, qP or F_v/F_m, thus suggesting that the efficiency of PSII activity and therefore linear electron flow were not affected by the absence of cytochrome c_{6A}. Uex55 and Uex87 did show a trend of lower calculated NPQ than cw15 (p=0.047 and 0.131 respectively), although this trend was not found with Uex78 which had a similar value of NPQ to cw15 (p=0.441). This meant that the cytochrome c_{6A} knockout line Uex55 had a significantly lower NPQ than cw15 upon SL exposure, and the cytochrome c_{6A} knockdown lines had a less clear NPQ relative to cw15. As the nature of Uex55 is the clearest, these data suggest that cytochrome c_{6A} may be involved in non-photochemical quenching.

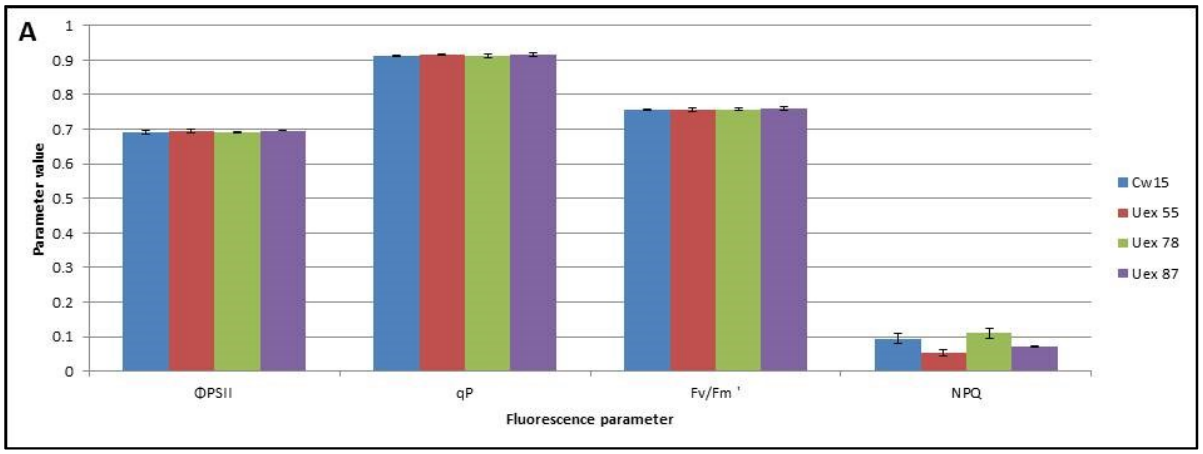


Figure 4.18: Calculated fluorescence parameters for the individual cytochrome c_{6A} knockout and knockdown lines and cw15. Error bars represent the standard error from three biological replicates.

4.2.5.2. Relaxation analysis was performed to analyse non-photochemical quenching further

In order to study the effect of cytochrome c_{6A} on NPQ further, relaxation analysis was performed on the Uex mutant lines and cw15. The cultures were exposed to high actinic light (500 µE m^{-2} s^{-1}) for 20 minutes before returning to dark conditions, with saturating pulses used to measure F_m' values every 5 minutes for 50 minutes (figure 4.19). By estimating the trajectory of fluorescence increase in the darkness, a theoretical value of how much quenching occurred immediately after the actinic light was switched off can be calculated (F_m^r; figure 4.20). As this value would only measure slow-relaxing quenching, it in turn can be used to calculate fast-acting and slow-acting NPQ (table 4.3), which can be attributed to qE and photoinhibition respectively (Maxwell and Johnson, 2000).

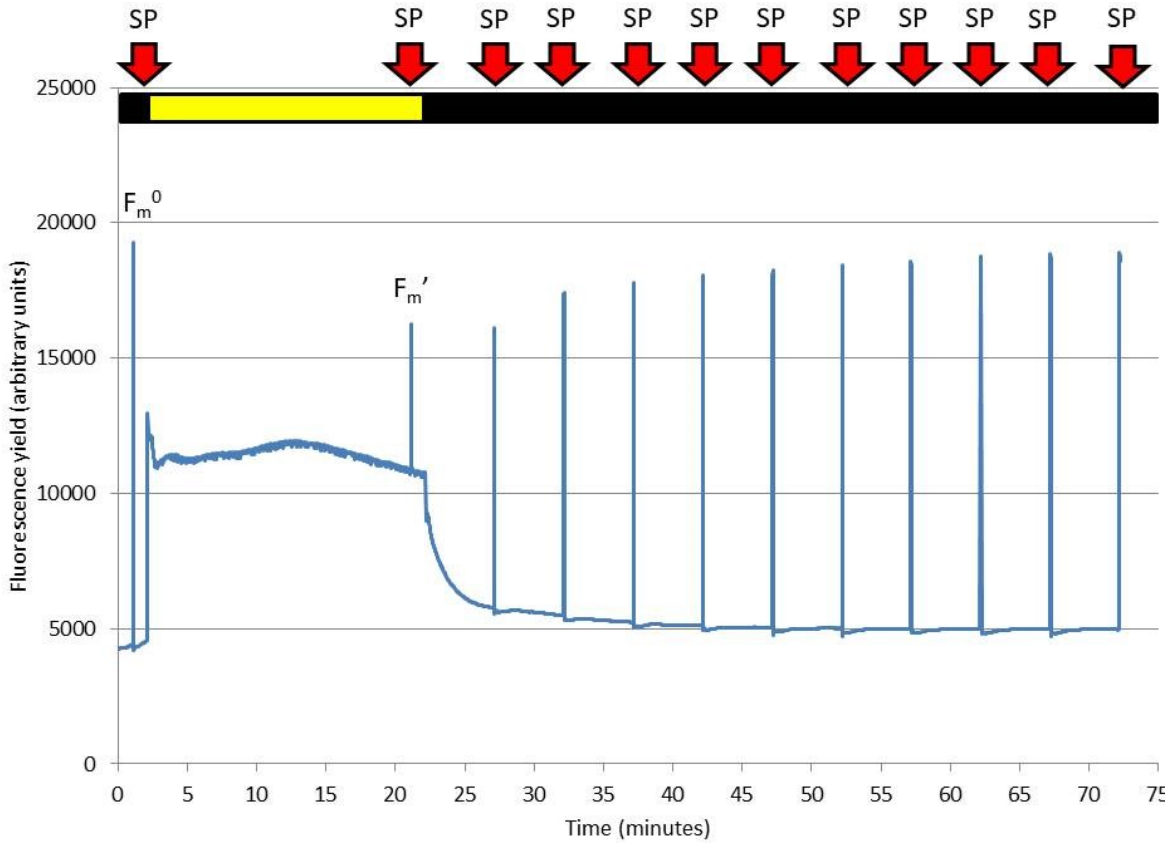


Figure 4.19: Trace taken from relaxation analysis light regime. After 10 min dark adaptation a measuring light (1.65 µE m^{-2} s^{-1}, black box) was turned on and a fluorescence trace was recorded. After a minute a saturating pulse (SP) was applied, giving F_m^0. After a further minute an actinic light of 500 µE m^{-2} s^{-1} (yellow box) was applied for 20 min. With 1 min remaining another SP was applied, giving F_m'. After the actinic light was switched off, SPs were applied every 5 min for 50 min. The fluorescence levels given in the post-actinic period were plotted on a log graph against time and a zero value extrapolated, named F_m^r. All SPs were of an intensity of 6172 µE m^{-2} s^{-1} for 800 ms, and the light used throughout was of 470 nm wavelength. Adapted from Walters and Horton (1991).

Non-Photochemical Quenching parameter	Formula
Slow NPQ (NPQ_S)	$\frac{(F_m0 - F_mr)}{F_mr}$
Fast NPQ (NPQ_F)	$\frac{F_m0}{F'_m} - \frac{F_m0}{F_mr}$

Table 4.3: Formulae for calculating the non-photochemical quenching parameters from the fluorescence levels measured or estimated in the trace in figure 4.19. Taken from Maxwell and Johnson (2000).

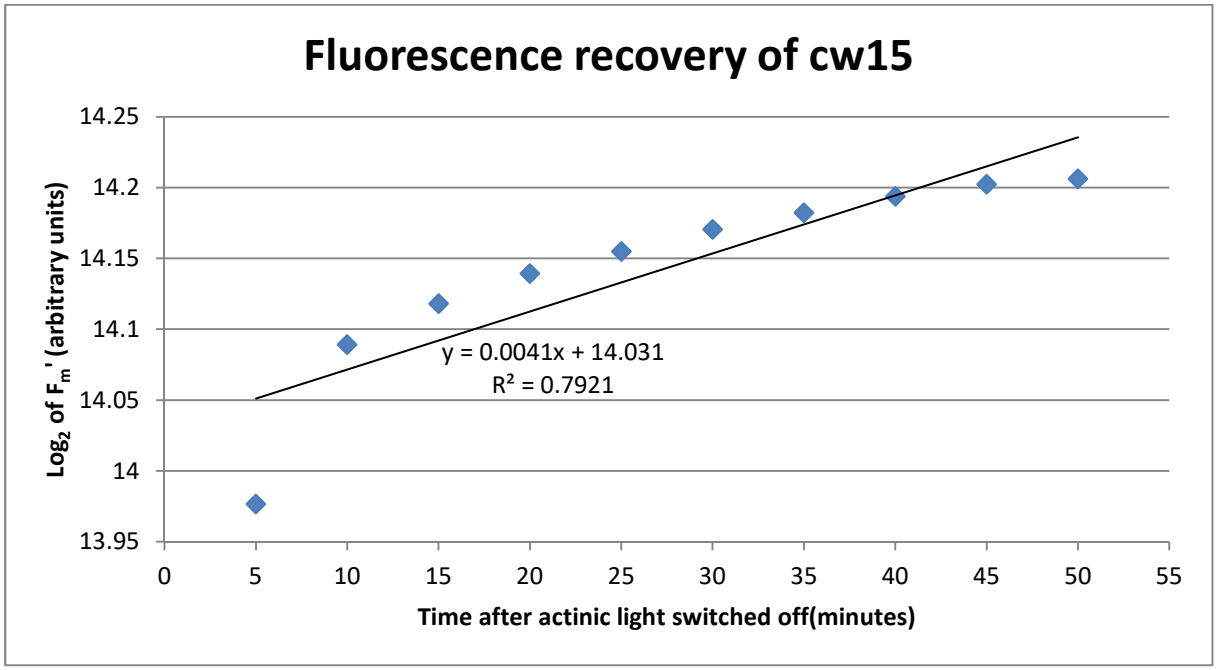


Figure 4.20: Graph plotting log_2 of the F_m' readings after exposure to actinic light readings over 50 minutes for cw15. A trendline has been drawn through these values to estimate the theoretical value of F_m immediately after the actinic light was switched off, in the absence of fast-relaxing quenching (F_m^r). The formula of the trendline has been given to show the potential zero value ($2^{14.031}$ = 16,740), along with the R^2 value.

NPQs and NPQf were calculated for each cytochrome c_{6A} Uex mutant line and cw15 (figure 4.21). Uex55 demonstrated a significantly lower NPQs than cw15 ($p<0.001$), but a similar value for NPQf ($p=0.469$). Uex78 and Uex87 however demonstrated a similar value of NPQs ($p=0.813$ and 0.155 respectively) and a significantly higher NPQf than cw15 ($p<0.001$ and $p=0.021$ respectively). As Uex55 has been found to be the only true knockout out of the mutants, in the case of this

inconsistency the conclusion is a phenotype of diminished NPQs when cytochrome c_{6A} has been knocked out.

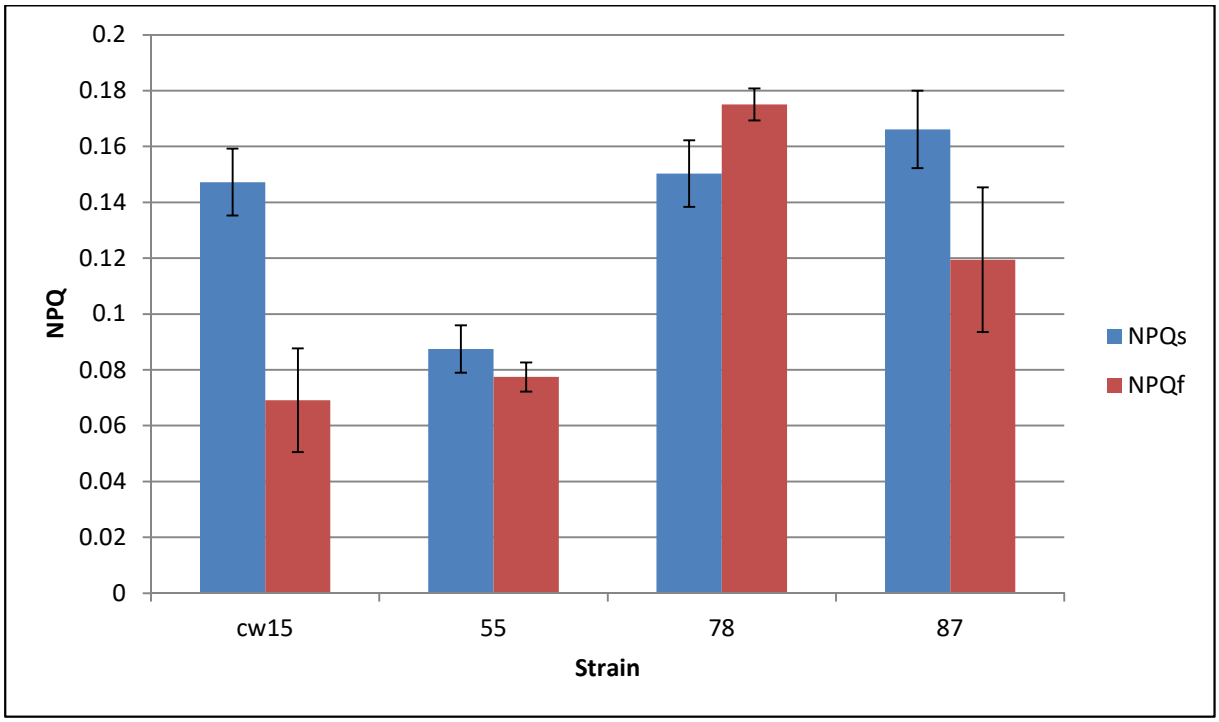


Figure 4.21: Calculated values of NPQs and NPQf for the cytochrome c_{6A} knockout and knockdown mutant lines and cw15. Values were calculated from the formulae in table 4.3. Error bars represent standard error from three biological replicates.

4.3. Discussion

4.3.1. Potential role for cytochrome c_{6A} in the high and fluctuating light stress response

4.3.1.1. Mixotrophic growth led to a potential growth phenotype upon high light stress

Under mixotrophic growth cytochrome c_{6A} mutant lines showed a trend of diminished growth when exposed to high light intensity. This would suggest that cytochrome c_{6A} could play a role in photoprotection, supported by the decreased NPQ observed in Uex55. Interestingly, the same trend was not observed under photoautotrophic growth, suggesting that cytochrome c_{6A} presence provides an advantage in high light stress only in the presence of an external source of organic carbon. *C. reinhardtii* grown in the presence of acetate has been shown to decrease PSII activity through state transitions and mild inhibition of water oxidation (Endo and Asada, 1996; Kühne *et al.*, 1999), therefore suggesting that cytochrome c_{6A} functions downstream of PSII in the electron

transport chain. Mixotrophic growth of *C. reinhardtii* has also been shown to decrease photoprotective measures around PSII (Polukhina *et al*., 2016). Cells grown mixotrophically have been found to have increased respiration and therefore CO_2 concentration in the cell (Fett and Coleman, 1994), which in turn leads to a decrease in LHCSR3 expression and NPQ (Peers *et al.*, 2009). Therefore this may explain why a potential phenotype was seen in TAP medium but not TP. If cytochrome c_{6A} is constitutively expressed at a low level, it could provide photoprotection (for example through the safety valve mechanism) at a similar constant low level. Then should other photoprotective measures be decreased under conditions of high CO_2 or carbon availability, cytochrome c_{6A} photoprotection could be observed. Also, if the presence of acetate is providing a carbon source such that CO_2 fixation is not essential, this latter process may be downregulated. Therefore if there is high light and there is a decrease in CO_2 fixation as a sink for NADPH generated from PSI, this might cause increased stress that cytochrome c_{6A} can help alleviate.

4.3.1.2. Mixotrophic growth led to a potential growth phenotype upon singlet oxygen stress

Under mixotrophic conditions the cytochrome c_{6A} knockout or knockdown lines also showed a trend of diminished growth upon RB exposure, both on solid medium and subsequently on liquid medium. The shape of the growth curves upon RB exposure suggested a longer lag phase for the mutant lines lacking cytochrome c_{6A}. One suggestion could be that the mutants were subject to a delay in response to singlet oxygen stress, for example in a signalling pathway leading to a longer term protective response, such as gene expression change. This could suggest a role for cytochrome c_{6A} in detection or signal transduction in response to singlet oxygen. Another explanation could be that the time taken to quench the singlet oxygen produced and repair PSII was longer in the absence of cytochrome c_{6A}, suggesting a potential role for cytochrome c_{6A} in singlet oxygen quenching or PSII repair.

The growth phenotype was again seen only with TAP medium and not in TP, which in a similar fashion to high light stress could be due to a decrease in photoprotective machinery in the TAP medium, particularly LHCSR3, leading to lower quenching of singlet oxygen produced. Acetate metabolism can provide reducing equivalents that can reduce plastoquinone. This would bypass PSII and potentially lead to retardation of quenching mechanisms, meaning that cells grown mixotrophically would be less prepared for singlet oxygen stress. This is supported by the observation that exposure to a small amount of RB allows for stronger protection against future exposure to higher amounts of singlet oxygen (Ledford *et al.* 2007). Again if cytochrome c_{6A} were

constitutively expressed regardless of acetate presence then it could provide photoprotection when the other mechanisms are lacking. Alternatively, as with the high light growth with acetate present, the decrease in CO_2 fixation would result in a decrease in sinks for NADPH produced at PSI, therefore meaning that under light there may be a higher incidence of singlet oxygen produced at PSII. This could combine with the exogenous singlet oxygen being provided by RB to make the cells more stressed without cytochrome c_{6A} photoprotective measures,

4.3.1.3. A growth phenotype was observed with fluctuating light and dark to light transitions

Cytochrome c_{6A} mutant lines showed a general trend of diminished growth under both fluctuating light and constant light after a period of darkness. This could suggest a role of cytochrome c_{6A} in short term responses to light stress, supported by the decreased NPQ observed in Uex55 relative to cw15. NPQ has been shown to increase during fluctuating light stress in *C. reinhardtti* (Janssen *et al.*, 2000), although the light fluctuations used were much more rapid than was used here. In fact, a criticism of the work described here could be that the fluctuating light scheme was actually diurnal and did not reflect the kind of rapid fluctuating light that NPQ would deal with. A mutant line of the cyanobacterium *Synechocystis* sp. PCC 603 lacking the terminal oxidases cytochrome *c* oxidase and quinol oxidase was observed to be diminished in growth following a 12 h on 12 h off light-dark cycle like the one used in this study (Lea-Smith *et al*., 2013). The work of Lea-Smith *et al.* shows that alternative electron pathways are essential for survival in fluctuating light in photosynthetic organisms over this diurnal cycle. Thus the apparent growth defect in fluctuating light seen here would be consistent with cytochrome c_{6A} being involved in an alternative electron pathway in *C. reinhardtii* that can provide short-term protection.

In a similar fashion to growth in the presence of acetate, growth in periods of darkness could lead to a decrease in photoprotective mechanisms, leading to higher sensitivity to stress when the light returns. If cytochrome c_{6A} were constitutively expressed and either provided an alternative route for electron flow through a 'safety valve', or was involved in sensing and initiating the response to light intensity, it could explain why moving from a period of darkness to a period of light would cause diminished growth in the Uex mutant lines.

4.3.1.4. Photoautotrophic growth led to a potential growth phenotype upon herbicide exposure

Under photoautotrophic conditions, cytochrome c_{6A} knockout or knockdown lines appeared to be impaired in growth upon exposure to DBMIB and DCMU. Both herbicides are plastoquinone analogues which can bind competitively to the Q_A^- pocket at PSII, but only DBMIB has been observed to inhibit cytochrome b_6f competitively at the Q_0 site (Chain and Malkin, 1979). Therefore the observation that both herbicides have a stronger effect on cytochrome c_{6A} knockout or knockdown lines suggests that cytochrome c_{6A} activity is related to stress at PSII, either directly or indirectly. A direct action could be to function in response to stress, such as increased ROS production caused by blockage in the linear electron transport chain, whereas an indirect action could be to allow an alternative path of electron transport to increase the efficiency of linear electron flow to help make up for the electron flow blocked by DBMIB and DCMU (safety valve).

Unlike the potential growth phenotypes observed for high light stress and singlet oxygen stress, herbicide treatment appeared to have a stronger effect on cytochrome c_{6A} knockout and knockdown lines in TP medium than in TAP medium. As LHCSR3 and photoprotective mechanisms have been observed to be increased under photoautotrophic growth relative to mixotrophic growth (Polukhina *et al*., 2016), the advantage of cytochrome c_{6A} presence when PSII oxidation is blocked could be more to do with increased electron flow through the electron transport chain than photoprotective measures such as singlet oxygen quenching. Also it has been observed that PSII activity was inhibited slightly under mixotrophic growth (Endo and Asada, 1996), therefore a more active PSII under photoautotrophic growth would lead to higher ROS production when blocked by DBMIB or DCMU, amplifying the photoprotective effect of cytochrome c_{6A}.

4.3.1.5. A majority of the growth curve trends observed were not statistically significant

A large caveat that needs to be noted when observing the trends in growth phenotype is that apart from in the mixotrophic high light conditions and with DBMIB in photoautotrophic conditions, no change in calculated growth rate for the cytochrome c_{6A} knockout or knockdown lines was found to be significant. This combined with very large error bars in many of the plots produced means that many of these observations are to be recognised as trends that require further repetition to determine reliability. The observations focussed on in the results and discussion as 'trends' were done so based on the consensus in change amongst the three Uex mutant lines, which suggests some level of reliability. One such reason for the subtle changes observed could be because

cytochrome c_{6A} expression is generally low (Nimmo, 2011), and therefore removing cytochrome c_{6A} may have only a relatively small effect proportionally.

4.3.2. Variation among the mutant lines

It can also be noted that in some studies there was inconsistency in growth between different the cytochrome c_{6A} mutant lines. This can be observed in the low light, RB TP and herbicide TAP growth curves in particular, as well as in the chlorophyll fluorescence measurements. This could be due to the genetic difference between mutant lines, as Uex55 was found to be the only strain that had the desired deletion after CRISPR-Cp1 treatment. Except for the herbicide growth curves in TP media, there was no consistent correlation between the mutant phenotype and the severity of the mutation, whereas it might have been expected that the fully knocked out Uex55 would show the largest change and the potential knockdown mutant lines Uex78 and Uex87 would show a smaller change. This underlines the need for further replicates.

4.3.3. Removing cytochrome c_{6A} did not affect photochemical quenching

There was no noticeable difference observed in ΦPSII, qP or F_v/F_m for the cytochrome c_{6A} knockout and knockdown mutants compared to cw15. This indicates that the rate of charge separation, photochemistry, and therefore the rate of linear electron flow were unaltered by the cytochrome c_{6A} mutations. This could be seen as evidence against a 'safety valve' model of cytochrome c_{6A} function whereby the protein provides an alternative route for electron flow from plastoquinol to plastocyanin or PSI. According to this model, providing an alternative route through cytochrome c_{6A} would be predicted to increase the rate of electron flow from PSII and therefore the efficiency of charge separation. However, these parameters were measured with an actinic light considered to be standard for *C. reinhardtii* growth (40 µE m^{-2} s^{-1}) and cytochrome c_{6A} levels are likely to be low in this condition. If cytochrome c_{6A} acts as a safety valve only when high light stress necessitates alternative electron flow, one might not see a significant effect on photosynthetic efficiency under standard light. Nevertheless, it appears cytochrome c_{6A} does not function to improve photosynthetic efficiency under standard light conditions.

4.3.4. A phenotype of diminished slow-relaxing NPQ can be observed when cytochrome c_{6A} is knocked out

Based on the fluorescence parameters calculated in the first experiment (figure 4.18), there was a trend for NPQ to be lower in most of the Uex lines compared to cw15. This would suggest a potential role for cytochrome c_{6A} in the heat dissipation of high energy state chlorophyll upon illumination. However, in the second relaxation analysis (figure 4.21) only Uex55 demonstrated slower NPQs, with Uex78 and Uex87 showing similar values for slow-reacting NPQ. The measured NPQf for the cytochrome c_{6A} mutants was unaltered in Uex55, but significantly higher in Uex78 and 87. As Uex55 was found to be the only true knockout mutant line, the overall observation of these experiments was that knocking out cytochrome c_{6A} lead to a significant decrease in slow-relaxing NPQ.

Considering the results with cw15 and Uex55 alone, these data could suggest that cytochrome c_{6A} plays an important role in the slower-response NPQ, perhaps suggesting a signalling role to affect the regulation of the xanthophyll cycle, photoinhibition, PSII antennae size, or the downstream sinking of electrons onto CO_2. The lack of significant change in NPQf, however, would suggest that cytochrome c_{6A} does not play a major role in the qE response, which combined with the lack of change in photochemical quenching observed with cytochrome c_{6A} knocked out provides further evidence against the safety valve hypothesis.

4.3.5. Differing actinic light and the extrapolation of the relaxation curves need to be considered when interpreting the results

One parameter that may add a level of uncertainty to the results is the actinic light used in each run. The first experiment used 6 min exposure to 40 µE m^{-2} s^{-1} and the relaxation analysis used a 20 min exposure to 500 µE m^{-2} s^{-1}. This latter use of high light intensity was based on other NPQ experiments in *C. reinhardtii* (Bonente *et al.*, 2011; Bonente *et al.*, 2012). However, this makes a direct comparison between the two experiments difficult as the relaxation analysis would be considered high light in *C. reinhardtii*. The lower light intensity used in the first experiment could also generate lower values of NPQ, as under 40 µE m^{-2} s^{-1} it is unlikely that the cultures will be stressed enough to activate photoprotective measures fully. This is supported by the observation that the NPQ values measured in this study are notably lower than those measured in prior studies on *C. reinhardtii* (Niyogi *et al.*, 1997; Ballottari *et al.*, 2016).

Another aspect of the relaxation analysis that needs to be noted is the quality of the extrapolation used to determine F_m^r (figure 4.20). The $\log_2$ graphs used for the extrapolation were not perfectly

linear, with the first value often showing a much lower value than expected for a linear trend. The subsequent points demonstrated a more linear trend, but removing the first value to improve the fit of the trendline would be unethical. However, only building the trendline with the first three points of the $\log_2$ curve generated negative values for NPQf, which is impossible in reality. Therefore all the points of f_m' were included in the trendline fit, leading to a lower R^2 value for the line, but a realistic value of NPQf. As this was performed for all the strains tested equally, the data are still comparable.

4.3.6. Conclusions and future work

The cytochrome c_{6A} knockout and knockdown lines have shown a mild trend of diminished growth when exposed to stress caused by high or fluctuating light, singlet oxygen and PSII-disrupting herbicides. Therefore this indicates that cytochrome c_{6A} may play a role in photoprotection, possibly by being involved in the detection of stress and activation of a response, or by providing an alternative electron pathway that would help electron flow through the transport chain and decrease the bottleneck effect. There appears to be an interaction between mixotrophic growth and cytochrome c_{6A} activity, suggesting that cytochrome c_{6A} is not downregulated (or perhaps even upregulated) in the presence of acetate, and can continue to provide photoprotection even when other photoacclimation machinery is not there and when CO_2 is not as strongly available as an electron sink for the PETC. Combining the growth curves and chlorophyll fluorescence measurements have also generated an argument for a role of cytochrome c_{6A}, directly or indirectly, in NPQ.

These growth curves provide the platform for many more experiments to explore this potential interaction of cytochrome c_{6A} with photosynthetic stress. Firstly repetition of these growth curves, particularly ones with high error, will show whether the trends observed are significant. Trying different length periods of fluctuating light, especially periods in the magnitude of seconds to minutes, may provide further insight into the role of cytochrome c_{6A} in dealing with rapid changes in intensity. Finally, an assumption for a number of the conclusions regarding cytochrome c_{6A} activity in mixotrophic growth is that cytochrome c_{6A} expression is unaffected or even increases when the cell is exposed to acetate. Therefore expression measurements of cytochrome c_{6A} when grown mixotrophically and photoautotrophically could provide more evidence behind cytochrome c_{6A} activity when acetate is available. Gene expression changes of cytochrome c_{6A} under photosynthetic stress could also provide insight into its function.

The chlorophyll fluorescence experiments in this study have ruled out a direct link between cytochrome c_{6A} and photochemical quenching and efficiency, and have found a potential involvement of cytochrome c_{6A} with slow-relaxing NPQ. To analyse further the potential effect of light intensity on cytochrome c_{6A} interaction with NPQ, both forms of experiment should be performed on the cytochrome c_{6A} knockout line Uex55 and cw15 after growth at varying light intensities. This should give more information on how cytochrome c_{6A} might assist or hinder the longer-term NPQ mechanisms, as well as whether photosynthetic efficiency under high light stress is affected by cytochrome c_{6A}. Both experiments could also be repeated using different intensities of actinic light. This should help in understanding the differences observed in the current study, whilst also providing more insight into cytochrome c_{6A} interaction with NPQ under different stress conditions.

5. Effects of acidic pH on cytochrome c_{6A} structure

5.1 Introduction

5.1.1. A decrease in pH in the thylakoid lumen activates the initial NPQ response

The rate of proton pumping across the thylakoid membrane, and therefore the pH of the thylakoid lumen, will fluctuate with changes in light intensity. During a period of high light intensity the availability of ADP and P_i becomes limiting, therefore the rate of ATP synthesis slows and the pH of the lumen drops due to the consequent retardation of proton gradient dissipation (Erickson *et al*., 2015). The optimal luminal pH for photosynthetic activity is thought to be between 6.8 and 7.5 (Kramer *et al.* 1999), and under high light the pH can fall as far as 4.0 (Jagendorf and Uribe, 1966; Krieger and Weiss, 1993). Therefore low pH in the lumen can be seen as an initial indicator of high light stress.

A lowered luminal pH leads to activation of early NPQ. This has been shown by the introduction of the H^+ ionophore nigericin to *C. reinhardtii* cells before or during high light exposure, which prevents or reverses NPQ respectively (Niyogi *et al.*, 1997a). Rapidly reversible NPQ (qE) in *C. reinhardtii* has two main mechanisms: through zeaxanthin and through a lutein cation (Amanarth *et al.*, 2012). Zeaxanthin has been found to bind the chlorophyll-binding protein LHCSR3, and *in vitro* studies have shown that this complex can form a carotenoid radical cation capable of rapidly quenching excited chlorophyll molecules (Bonente *et al.*, 2010). LHCSR3 has also been shown to be able to respond to pH change, with truncated forms of the protein lacking the C-terminal sequence being unable to quench excited chlorophyll in response to low pH (Ligouri *et al.*, 2013). The study therefore suggests that a low pH alters the conformation of the LHCSR3 C-terminal region, which allows LHCSR3 to act in qE.

The lutein cation mechanism was found using a mutant strain (*lor1*) with a non-functional lycopene-ε -cyclase (Anwaruzzaman *et al.,* 2004). This impaired the synthesis of α-carotene derived xanthophylls, and was shown to have diminished pH-induced NPQ (Niyogi *et al.*, 1997b). It was also shown that a double mutant of *lor1*, *npq1*, which lacks the ability to synthesise either zeaxanthin or lutein, completely lost qE (figure 5.1; Niyogi *et al.*, 1997b), therefore suggesting that the two systems separately contribute to chlorophyll quenching in response to low pH.

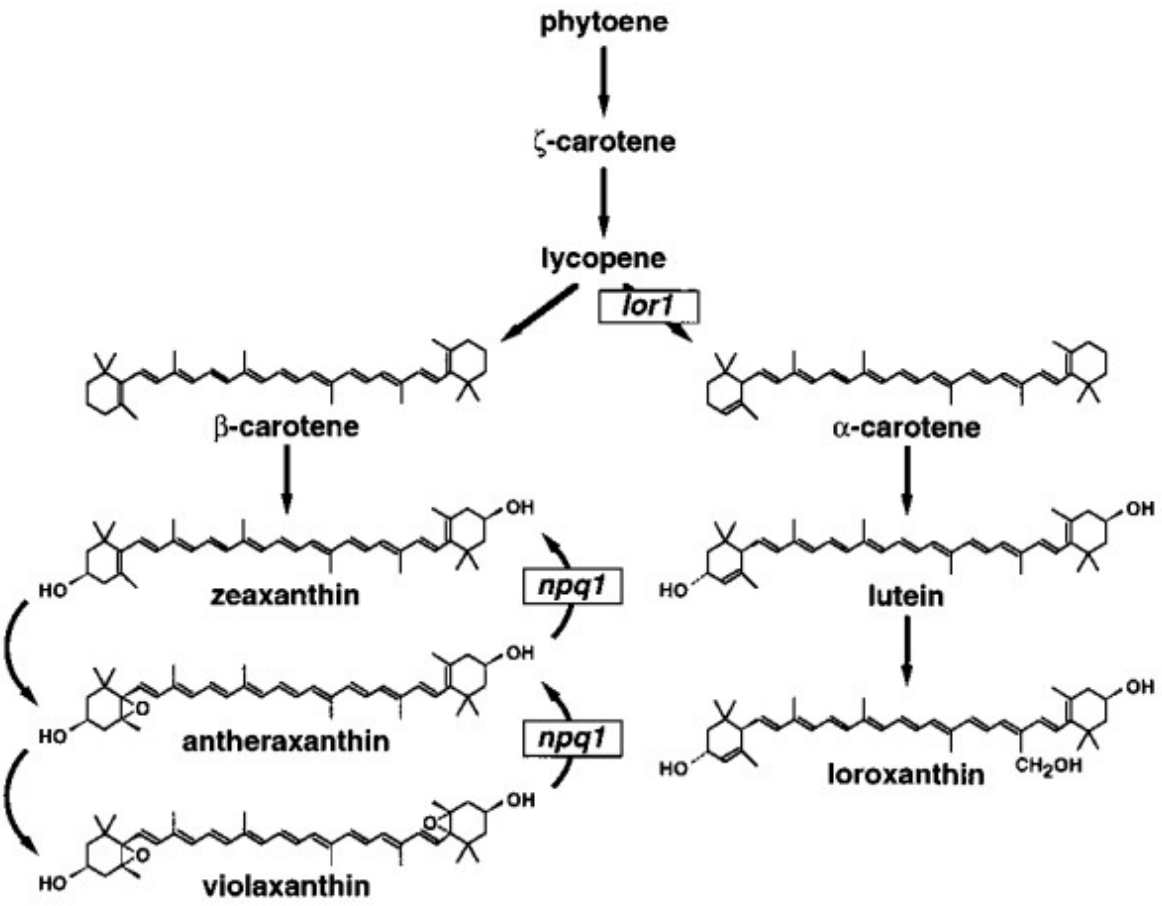


Figure 5.1: Biosynthetic pathway for the carotenoids zeaxanthin and lutein in C. reinhardtii. The reactions disrupted by the mutant lines lor1 and npq1 have been indicated. Adapted from Niyogi et al. (1997b).

5.1.2. Potential for cytochrome c_{6A} to respond to pH

Cytochrome c_{6A} has been found to be located in the thylakoid lumen of *C. reinhardtii* using fluorescent tagging (Nimmo, 2011). It has already been shown with X-ray crystallography that the oxidation state of the cytochrome c_{6A} haem prosthetic group does not alter the structure significantly (Marcaida *et al.*, 2006). However, the pH of the environment could affect cytochrome c_{6A} structure in a similar way to LHCSR3, converting the protein between an active or inactive conformation. Furthermore, NMR performed on cytochrome c_{6A} has shown potentially interesting NMR shifts of certain amino acids when the pH was increased from 5.86 to 6.45 (Nimmo, 2011). The amino acids with the largest shift were found to be in a loop region between two alpha helices close to one of the propionate groups of the haem cofactor (V43, D44 and T45; figure 5.2), suggesting that

under pH change the relative positions of the alpha helices in cytochrome c_{6A} shift. Therefore structural studies on cytochrome c_{6A} at more acidic pHs would reveal more potential conformational changes, which can indicate a function in response to high light stress.

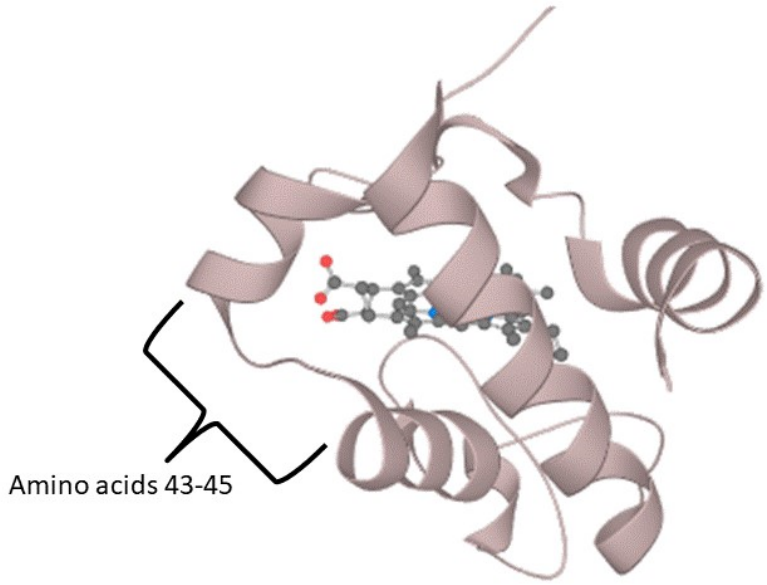


Figure 5.2: Protein ribbon structure of A. thaliana cytochrome c_{6A} with amino acids 43,44 and 45 indicated. Spiral ribbons represent alpha helices and long thin ribbons represent loop regions. The haem prosthetic group is represented by the ball and stick model, with black, red and blue balls representing carbon, oxygen and iron atoms respectively. Created using structure viewer on Uniprot using X-ray crystal structure from Marcaida et al. (2006).

Circular dichroism (CD) is a spectroscopic method of determining broad structural changes in proteins, particularly secondary structures (Chen *et al.*, 1972; Greenfield, 2006). At UV wavelengths, different secondary structures display uneven absorption of left and right circulatory plane polarised light, therefore giving distinct ellipticity spectra that can be compared to spectra of other proteins to determine the relative amount of alpha helices, beta strands and loop regions (figure 5.3). Alpha helices tend to demonstrate two negative peaks at around 222 and 208 nm with a positive peak at 192 nm. Proteins containing antiparallel beta sheets on the other hand demonstrate a positive peak at 195 nm, with a negative peak at around 218 nm. Disordered proteins can have a negative peak around 195 nm, but demonstrate poor ellipticity elsewhere. Therefore CD would give an idea of any large changes in cytochrome c_{6A} structure upon exposure to acidic pH, with a particular focus on secondary structure changes and evidence of denaturation.

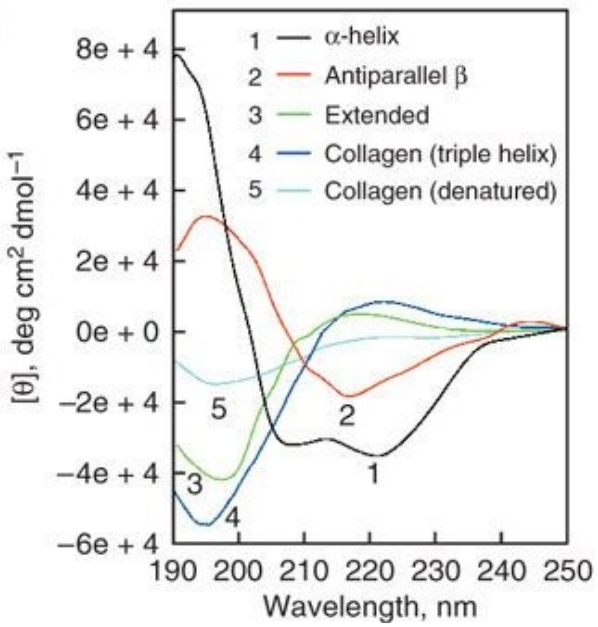


Figure 5.3: Circular dichroism spectra for poly-(L)-lysine peptides demonstrating alpha helical (1), beta sheet (2) or extended (3, no specific secondary structure) conformations. Results from collagen in its native form (4) and denatured (5) are also given to show how denaturing affects ellipticity spectra. Adapted from Greenfield, 2006.

5.1.3. Aim of the study

- To determine any broad signs of cytochrome c_{6A} structural change with decreasing pH.

5.2. Results

5.2.1. Expression and purification of cytochrome c_{6A} in *E. coli*

E. coli strain GM119 was transformed with the plasmid pATc6A encoding *A. thaliana* cytochrome c_{6A} and with the plasmid pEC86 encoding a haem maturation cassette (Arslan *et al.*, 1998), hereby named *E. coli* BS1_c6A. The transformants were selected through growth in the presence of chloramphenicol selecting for 25 µg/mL and ampicillin selecting for 100 µg/mL and confirmed through plasmid extraction of a sample of the transformed culture and digestion with the enzyme *Pvu*I (figure 5.4). *E. coli* transformed with peEC86 would show bands of 7244 bp and 4509 bp, and cells transformed with pATc6A would show bands of 895 bp and 2416 bp. The presence of all four fragments in the *Pvu*I digestion for cells isolated from a single colony showed a successful double transformation.

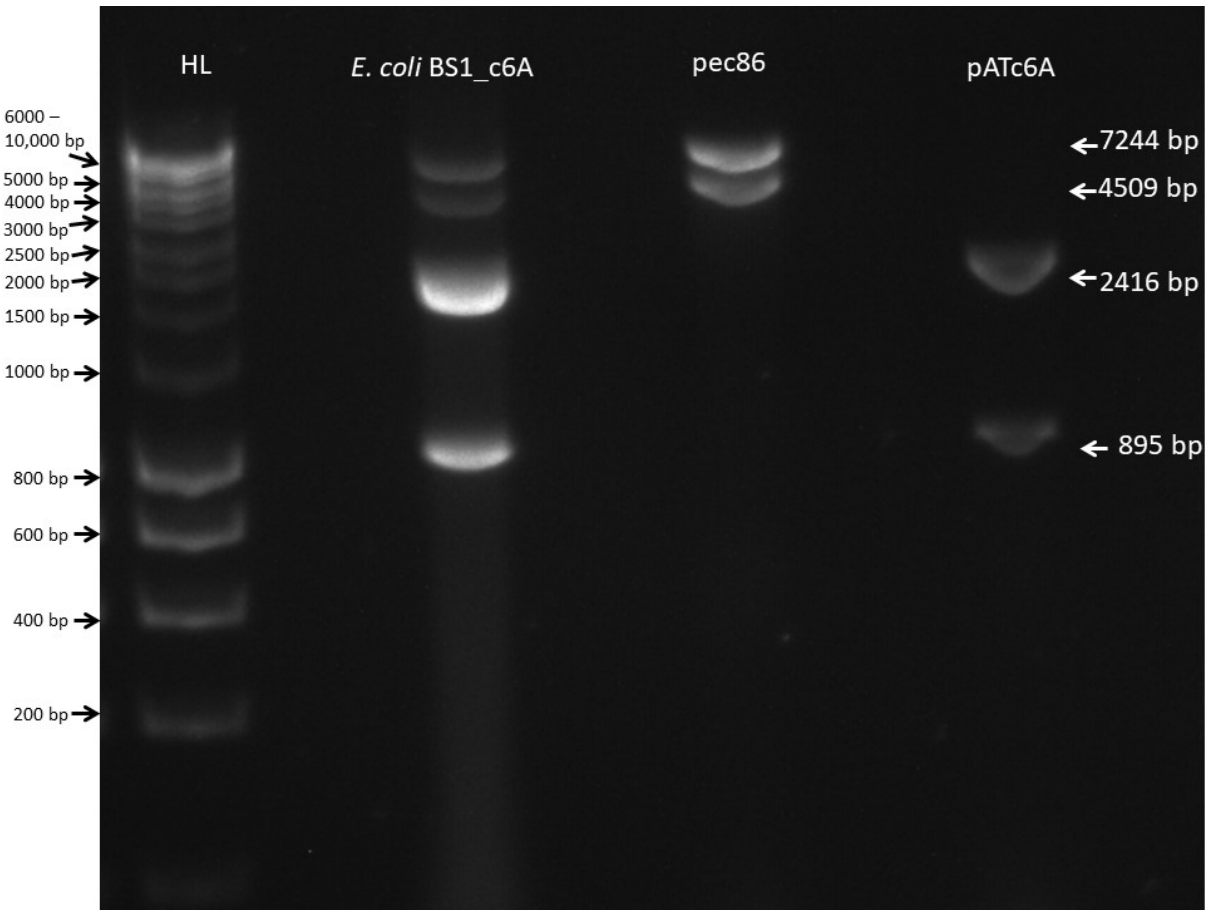


Figure 5.4: Confirmation of E. coli transformation with both pEC86 and pATc6A through pvuI digestion. The first lane contains Hyperladder 1kb. The second lane contains DNA extracted from strain E. coli BS1_c6A, the third lane contains pure pec86 and the forth lane contains pure pATc6A. All of lanes 3-5 were digested with PvuI.

An *E.* coli strain *E. coli* BS1_c6A that was confirmed to contain both plasmids was used to inoculate larger cultures of 1.7 L and grown for ~80 h. Cell pellets of the culture were observed to have a pink colouring (figure 5.5A) due to the incorporated haem prosthetic group of cytochrome c_{6A}, indicating successful expression.

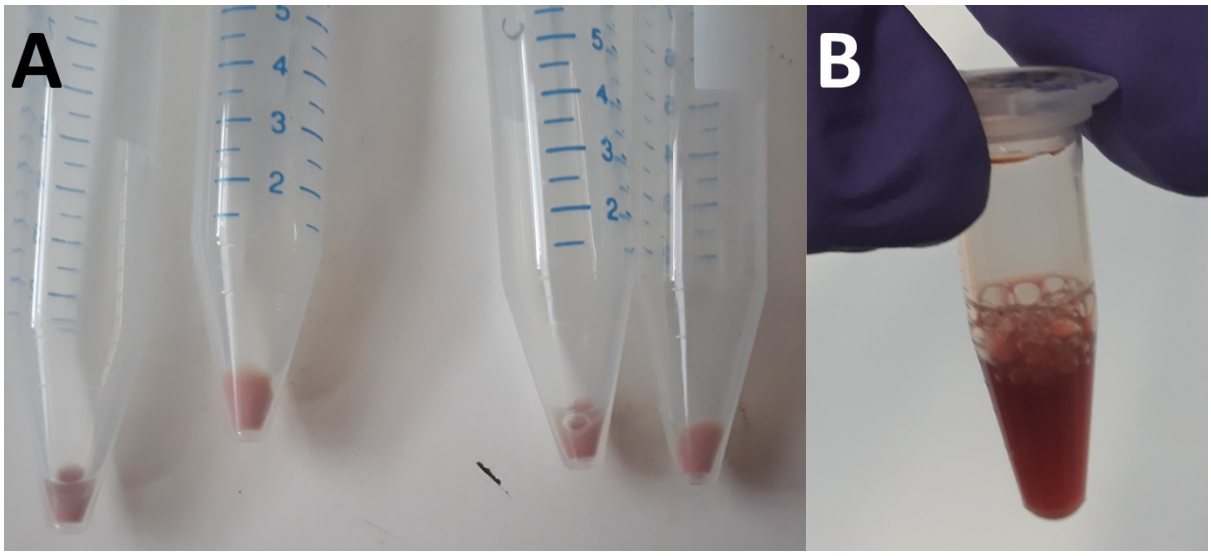
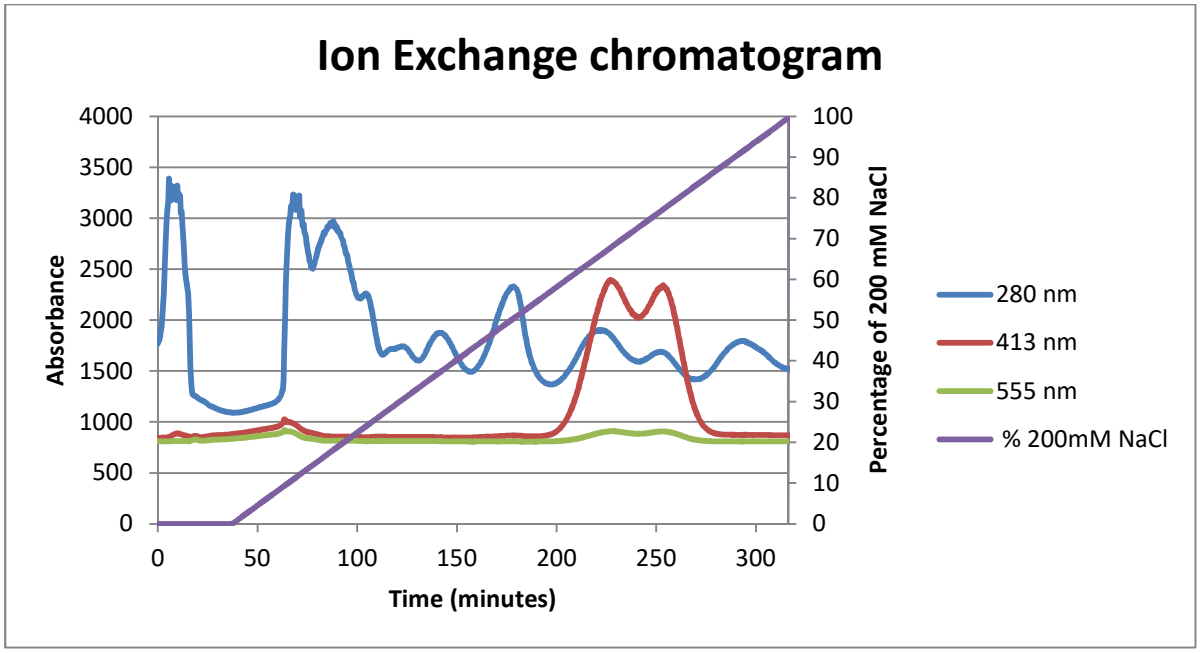


Figure 5.5: A) Pink pellets from the centrifugation of E. coli BS1_c6A*. Each tube is from a separate culture grown from individual colonies on a transformed plate. The pink colour indicates successful expression of cytochrome* c_{6A}*. B) Purified cytochrome* c_{6A} *after both runs of chromatography and concentration.*

Purification was carried out as described in the Materials and Methods chapter, section 2.7.5. The protein was extracted from the *E. coli* BS1_c6A cultures, filtered and then loaded into an ion exchange column (figure 5.6). Cytochrome c_{6A} was detected by measuring absorbance at 413 and 555 nm wavelengths of the eluate, supported by observing a faint red colour to the fractions chosen. Absorption at 280 nm was also measured to show elution of all proteins. Unexpectedly, two peaks in absorbance of 413 and 555 nm were detected, with red protein eluting in two sets of fractions.



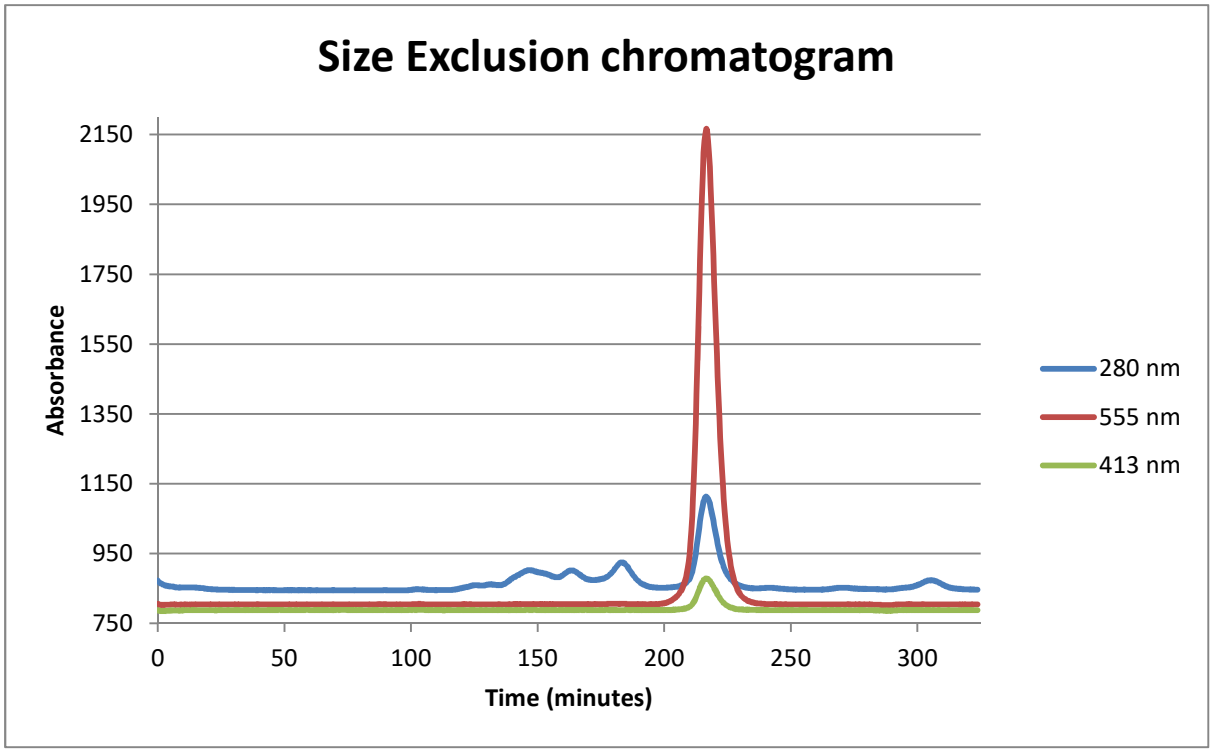


Figure 5.6: Chromatograms for FPLC runs through an ion exchange column (above) and a size exclusion column (below). Fractions were taken every 1.9 mL.

Both sets of pink fractions from the ion exchange were pooled, concentrated and loaded into a size exclusion column (figure 5.6). Again, absorbances at 280, 413 and 555 nm were measured, with a single large peak emerging for the latter two wavelengths after around 220 mL was passed through.

The fractions eluted at this point again had a faint red colour. They were further concentrated, with the red colour becoming more prominent (figure 5.5B).

5.2.2. Analysing the purity and integrity of purified cytochrome c_{6A}

Presence of other proteins in the purified cytochrome c_{6A} sample would drastically affect the CD measurements. Therefore the *E. coli* lysate, and samples of fractions collected after ion exchange and size exclusion protein were analysed by SDS PAGE (figure 5.7). A band of around 11.8 kDa was the only visible band after size exclusion, suggesting that the fraction was almost entirely cytochrome c_{6A}. The presence of other proteins in the fraction after ion exchange demonstrated the importance of running the *E. coli* lysate down both columns. A western blot of the same samples using an antibody against *A. thaliana* cytochrome c_{6A} confirmed the identity of the band (data not shown).

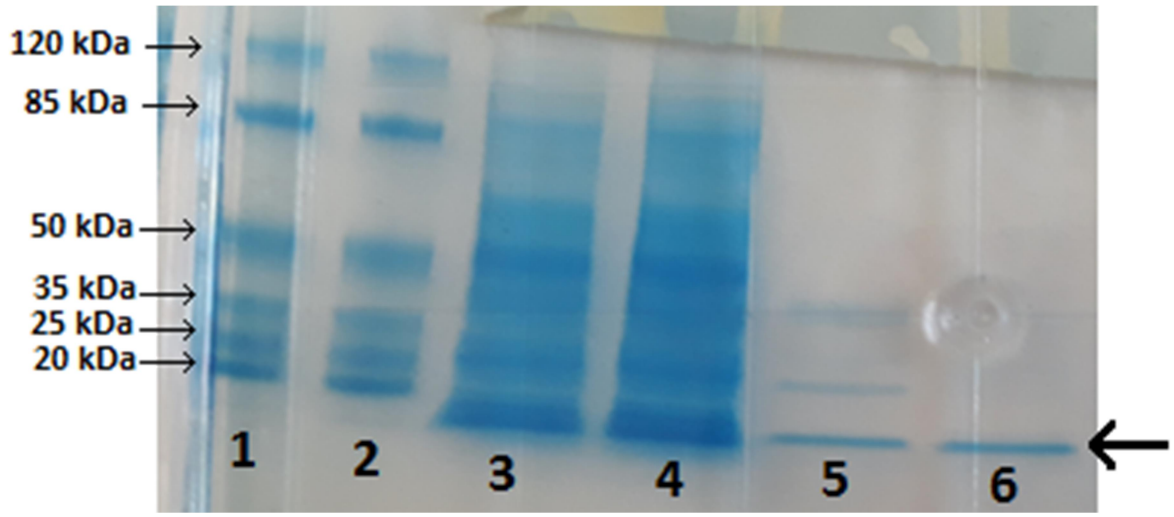


Figure 5.7: SDS-PAGE of the process of cytochrome c_{6A} purification from E. coli lysate (lanes 3 and 4, separate cultures). Lanes 1 and 2 contain Pierce prestained protein ladder, lane 5 contains the cytochrome c_{6A} fraction after ion exchange chromatography, and lane 6 contains the cytochrome c_{6A} fraction after size exclusion chromatography. The band representing cytochrome c_{6A} (around 11.8 kDa) is indicated by the arrow.

5.2.3. Circular dichroism showed cytochrome c_{6A} maintains structural integrity down to pH 2

Aliquots of the purified, concentrated cytochrome c_{6A} were transferred to buffers of pH 2, 3, 4, 5, 6 and 7. For pH 2, 3, 7 and 7 phosphate buffer was used, and for pH 5 and 6 acetate buffer was used.

The ellipticity of the cytochrome c_{6A} solutions was then measured through CD (figure 5.8). Each spectrum showed strong similarity with a typical alpha helical protein (figure 5.3), with a strong positive peak at around 192 nm and two negative peaks at 208 and 222 nm. All the spectra looked very similar in shape, suggesting little difference in the secondary structure and denaturation of cytochrome c_{6A} between pH 2 and 7. The most variation came from pH 4 and 5, where the former spectrum appeared to be shifted down, and the latter appeared to be shifted up, particularly between 200 and 225 nm. The spectrum between 190 and 200 nm at pH 5 was also highly variable, giving very large error bars. These differences in spectra may suggest mild structural differences occurring in cytochrome c_{6A} at pH 4 and 5 relative to pH 2, 3, 6 and 7.

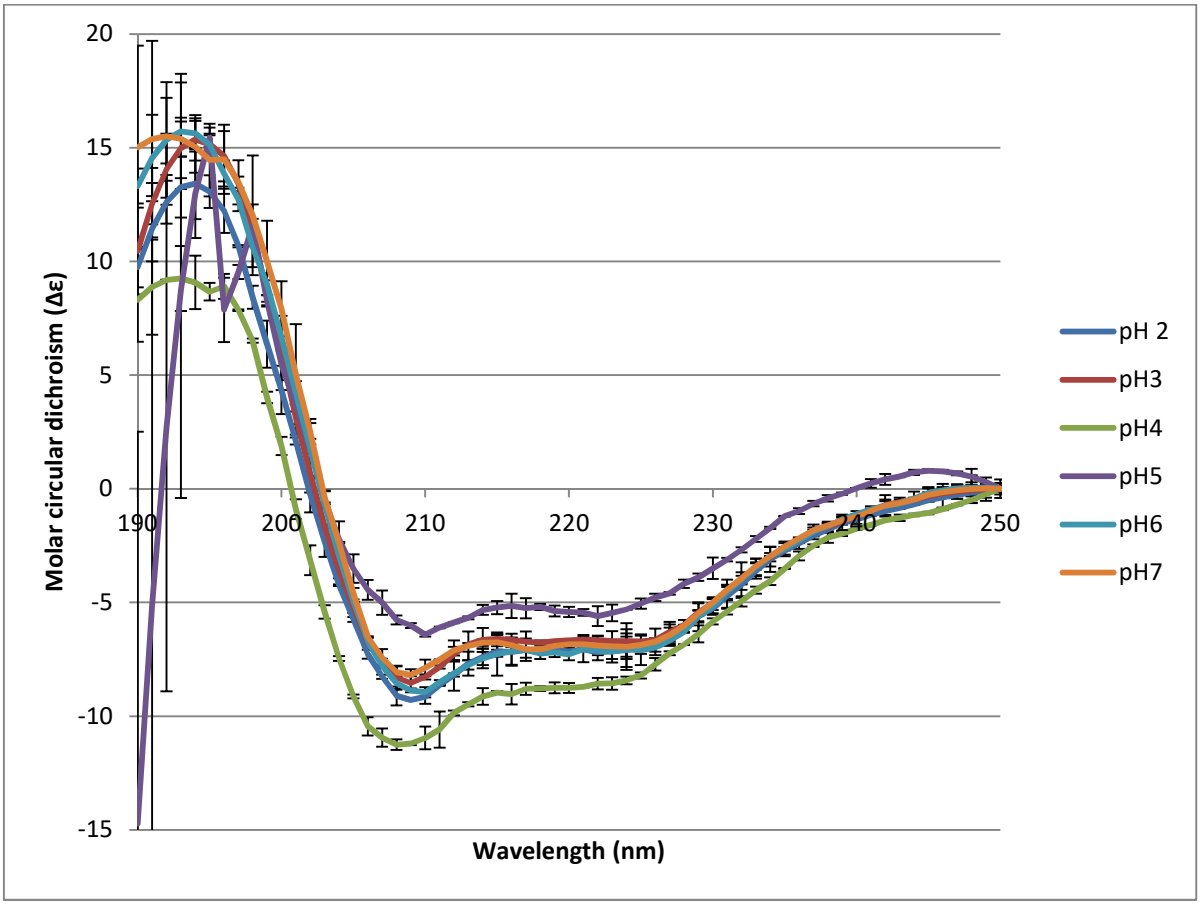


Figure 5.8: Circular dichroism spectra for cytochrome c_{6A} Exposed to different pH buffers. The spectra have been normalised to wavelength 250 nm. Error bars demonstrate standard error. The error bars at 190 and 191 nm for pH 5 has been cut off for easier visual, but extend to -31 and -24 respectively.

From these data the proportion of alpha helix relative to the rest of the protein was estimated by two different methods (figure 5.9). The ratio of ellipticity at 222 nm and 208 nm gives an estimate of how much of the protein is in alpha helix conformation relative to the looser loop structure, with a higher 222:208 ratio suggesting more alpha helix (Kallenbach *et al.*, 1996). This ratio value was found

to be very similar across most of the pHs, with the exception of pH 5 which demonstrated a much higher 222:208 ratio, suggesting that more of the protein was found to be in the alpha helix conformation than the other conditions. The alpha helix percentage prediction used the "calculate from ellipticity" algorithm from the Dicroprot software package (Deléage and Geourjon, 1993). Again, the values were relatively similar except in the cases of pH 4 and 5, which showed a slightly higher and lower percentage predicted respectively. This reflects the shifts in spectra seen in figure 5.8, with the more negative peaks resulting in a higher predicted percentage. It is worth noting that for pH 5, the two methods disagree drastically in prediction of alpha helix content. Overall it appears that cytochrome c_{6A} maintains structural integrity between pH 2 and 7.

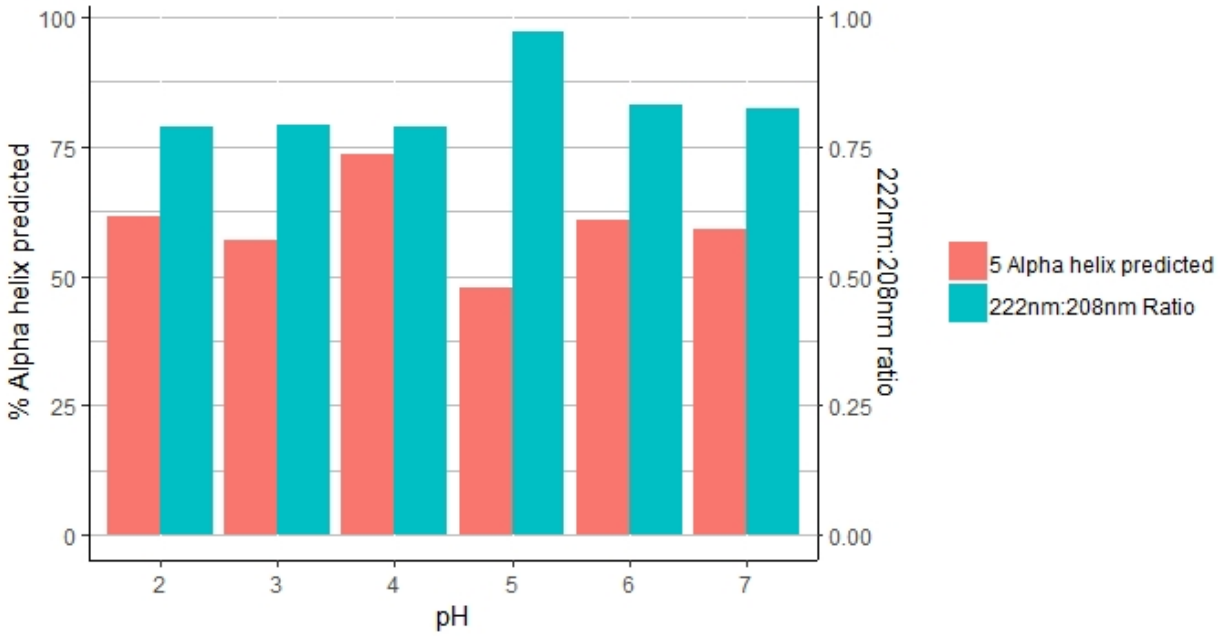


Figure 5.9: Predicted alpha helix percentage and 222nm:208nm ratios calculated for cytochrome c_{6A} at different pHs.

5.3. Discussion

5.3.1. Behaviour of cytochrome c_{6A} during purification

The double peak in absorbance at 555 and 413 nm measured during ion exchange chromatography raises concern that two forms of cytochrome c_{6A}, or even another red protein, was in the *E. coli* lysate. However, subsequent analysis of the purified protein, firstly through size exclusion and then through SDS-PAGE and western blot, showed the presence of only one protein of one size. Therefore it is more likely that a fault in the column or the detection measurements led to the cytochrome c_{6A}

eluting at two salt concentrations. However there is a possibility that a minor chemical modification such as reduction or oxidation of amino acid residues could have affected the charge properties but not the molecular weight, explaining the separation in ion exchange chromatography but not in size exclusion.

5.3.2. Comparison of cytochrome c_{6A} to other proteins found to change under varying pH

Circular dichroism has been used in the past to determine how secondary structures of proteins change upon exposure to differing pH (Su and Jirgensons, 1977; Mäkinen *et al*., 2016; Mantovani *et al.*, 2016). However, it is worth noting that in many such studies the structural change has been denaturation to varying extents, rather than more subtle conformational changes. A typical observation in such studies is a reduction in magnitude of the peaks corresponding to specific secondary structures, as shown in figure 5.3. For alpha helices in particular, a notable reduction in the 222:206 nm ellipticity ratio can be seen on denaturation (Su and Jirgensons, 1977), as the alpha helix conformations denature to form free loop structures. The changes observed in these studies are far more extreme than any change noted in this study, further suggesting that cytochrome c_{6A} remained structurally stable between pH 2 and 7.

5.3.3. Changes observed at pH 4 and 5 are most likely due to differences in buffer

The shifts in spectra and differences in predicted alpha helix proportions recorded for pH 4 and 5 could suggest structural changes for cytochrome c_{6A}. From the X-ray crystallography structure of cytochrome c_{6A} (Marcaida *et al.*, 2006) the alpha helix content can be estimated at around 54 %. This value is slightly lower than calculated for pH 2, 3, 6 and 7 by CD and therefore shows an overestimation of the alpha helix content in this study, possibly due to the high variance in ellipticity below 200 nm. The calculated alpha helix content for pH 4 is considerably higher than 54 %, and that for pH 5 was slightly lower. This could be significant as this pH range has been predicted to reflect the acidity of the thylakoid lumen upon high light stress (Krieger and Weiss, 1993), and therefore a conformational change in response to this drop in pH could be detected by cytochrome c_{6A} as a part of a function in qE. However, acetate buffers were used for the pH 4 and 5 conditions, and acetate absorbs in the far-UV range much more than monobasic or dibasic phosphate (Kelly *et al*., 2005). Therefore the use of different buffers could affect the relative absorption of the wavelengths used in this CD spectroscopy, making comparison between the acetate buffer pHs and the phosphate buffer pHs difficult. This, combined with the observation that the predicted alpha helix content of

cytochrome c_{6A} at pH 5 differs drastically depending on the method used (figure 5.9) suggests that differences in the buffer content may be responsible for the changes observed when comparing pH 4 and 5 to the other pH spectra. This interpretation is also supported by the large error observed at the far UV wavelengths in the pH 5 spectrum. Therefore it would appear that secondary structure of cytochrome c_{6A} remains intact across pH 2-7. However, the study cannot rule out subtle changes in structure not detectable by CD.

5.3.4. Other potential effects of pH on cytochrome c_{6A}

Changes in pH in the thylakoid lumen could also affect cytochrome c_{6A} redox midpoint potential and electrostatic surface charge distribution. The interaction between cytochrome c_{6A} and plastocyanin and the redox midpoint potential of cytochrome c_{6A} were recorded at pH6 and 7 respectively (Marcaida *et al.*, 2006) and therefore also would not reflect the luminal conditions under high light stress. A decrease in pH could potentially increase the redox midpoint potential due to an increase in proton availability for reduction, and therefore affect interaction partners of cytochrome c_{6A}. Similarly, the electrostatic surface charge distributions of cytochrome c_{6A} were based on a pH of 7, and therefore also do not reflect luminal conditions under high light. An increase in protons around cytochrome c_{6A} could make the overall charge distribution more positive, which may affect interactions with plastocyanin, PSI or other proteins in the lumen. Finally, despite the pH change not seeming to break the disulphide bridge of cytochrome c_{6A} in this study, pH changes could affect thiol redox chemistry as it has done for other luminal proteins upon light exposure (Buchanan and Luan, 2005). Therefore despite no significant structural changes being observed under differing pH, there are other ways that cytochrome c_{6A} could potentially 'detect' luminal pH.

5.3.5. Implications for cytochrome c_{6A} function in non-photochemical quenching

Cytochrome c_{6A} did not show drastic changes in secondary structure upon increasingly acidic pH. This suggests that cytochrome c_{6A} could remain functional even upon extreme changes in pH, which would be a very useful feature in both high light stress and more specifically qE. Cytochrome c_{6A} does not appear to undergo major conformational changes to become activated or inactivated, although smaller conformational changes cannot be ruled out. This suggests that cytochrome c_{6A} is unlikely to be changing structure drastically in response to luminal pH, but could still be involved in signalling by responding to luminal pH in a different way (see section 5.3.3). This could be electron transfer in a signalling pathway in response to the pH drop. Alternatively, if the redox midpoint potential of cytochrome c_{6A} allows oxidation of an over-reduced plastoquinone pool, the electrons

could be used in the redox chemistry used in carotenoid biosynthesis to increase photochemical quenchers under high light stress (Couso *et al.*, 2012).

5.3.6. Future investigation into cytochrome c_{6A} structure, response to pH and relation to NPQ

The results of this study provide a platform for potentially interesting future experiments connecting cytochrome c_{6A} structure to function. Circular dichroism is an effective tool for predicting secondary structure and determining denaturation of proteins, but does not show specific amino acid positions or tertiary structure detail. Therefore a higher resolution spectroscopy technic such as NMR could provide more detail on cytochrome c_{6A} structural changes upon pH shifts. Performing similar structural studies with a modified form of cytochrome c_{6A} without the LIP could also show if the insertion provides extra structural support at lower pH, which could provide a reason for its existence. The effect of lower pH on electrostatic charge distribution, redox midpoint potential and interactions with plastocyanin should also be investigated to determine any other mechanism for cytochrome c_{6A} to respond to luminal pH.

Further studies could also be performed to determine cytochrome c_{6A} involvement in NPQ, and specifically qE. Interaction studies such as pull-down assays could be performed to determine if cytochrome c_{6A} interacts with other proteins known to be involved in qE, such as LHCSR3 and the biosynthesis enzymes for carotenoids. The carotenoid content of *C. reinhardtiii* strains with cytochrome c_{6A} knocked out or overexpressed could show interactions between cytochrome c_{6A} and carotenoid synthesis. These further studies would bring more clarity on cytochrome c_{6A} involvement in NPQ.

6. Transcriptomic analysis of cytochrome c_{6A} knock out response to high light stress using RNA sequencing

6.1. Introduction

6.1.1. Previous studies have analysed transcriptomic regulation in response to ROS and high light stress in *Chlamydomonas reinhardtii*

6.1.1.1. The transcriptome response to ROS

RNA sequencing studies have been performed to analyse the effect of singlet oxygen and hydrogen peroxide on the *C. reinhardtii* transcriptome (Urzica *et al.*, 2012; Wakao *et al.*, 2014). Hydrogen peroxide is a ROS primarily formed upon high light exposure as a result of superoxide formation and dismutation at PSI (Patterson and Myers, 1973; Takahashi *et al.*, 1998). RNAseq was performed on *C. reinhardtii* exposed to 1 nM H_2O_2 for either 30 or 60 minutes as part of a study on the function and regulation of a GDP-L-galactose phosphorylase VTC2, which was found in the study to be a highly regulated enzyme in ascorbate biosynthesis (Urzica *et al.*, 2012). Transcripts whose levels increased following hydrogen peroxide exposure included those for enzymes involved in the glutathione ascorbate pathway such as ascorbate peroxidase (APX1) and glutathione reductase (GSH1), and superoxide dismutases such as MSD3. Singlet oxygen is a ROS formed primarily at PSII upon high light exposure, and induces a retrograde signalling pathway affecting nuclear transcription (Fischer *et al.*, 2007). In a study demonstrating the involvement of the transcription factor SAK1 in this singlet oxygen signalling pathway, an RNAseq experiment was performed on *C. reinhardtii* after 1 h exposure to 1 µM of the singlet oxygen producing chemical Rose Bengal (Wakao *et al.*, 2014). Genes showing a notable change in transcript levels upon singlet oxygen exposure included those encoding the glutathione peroxidase homologue GPXH, cyclopropane fatty acid synthases CFA1 and CFA2, and the haem-binding proteins SOUL1 and SOUL2. These genes could therefore be considered markers for transcriptional regulation in response to either ROS.

6.1.1.2. The transcriptome response to high light stress

A systems analysis has been performed on *C. reinhardtii* during exposure to increased light intensity (Mettler *et al.*, 2014). This study observed the effects of increased irradiance on metabolomics, proteomics, and transcriptomics, the last of which was performed through a microarray. Cultures in the study were grown under 41 µE m^{-2} s^{-1} before exposure to 145 µE m^{-2} s^{-1} and samples taken at 40,

120 and 240 minutes after the increase in light intensity. Transcripts noted in the study as increasing significantly in abundance upon this change in light intensity include the light harvesting complex related protein LHCSR1, carbon concentration proteins CCP2 and LCI1, and the nucleoredoxin NRX3. However, the increase in irradiance was only triple what was considered standard light, and the habitat of *C. reinhardtii* could go much higher than 150 µE m^{-2} s^{-1}. It is worth noting that throughout the literature there is no consensus on what should be considered "high light" for *C. reinhardtii*, with some papers using intensities as high as 2,500 µE m^{-2} s^{-1} (Fischer *et al.*, 2006; McKim *et al.*, 2006). Both isoforms of another light harvesting complex related protein LHCSR3, have been found to be strongly upregulated in response to high light exposure (Maruyama *et al.*, 2014), and thus both can also be considered potential markers for high light transcriptional change in *C. reinhardtii*.

If cytochrome c_{6A} is involved in high light stress as a safety valve response or a signalling intermediate, then knocking out cytochrome c_{6A} would be expected to have an effect on the transcriptome of *C. reinhardtii* on high light exposure. For example, should cytochrome c_{6A} act as a safety valve then a mutant line without it may show increased upregulation of light stress transcript upon high light exposure. On the other hand if cytochrome c_{6A} functions as a signalling protein then knocking out the gene may diminish transcriptional regulation upon increased irradiance. If any of these changes in transcriptome show extensive overlap with the studies previously performed on *C. reinhardtii*, then it will show a potential link between cytochrome c_{6A} and that stress response. Therefore in this study the transcriptomes of the knock out mutant line Uex55 and the background strain cw15 were analysed through RNA sequencing upon exposure to a ten-fold increase in irradiance (figure 6.1).

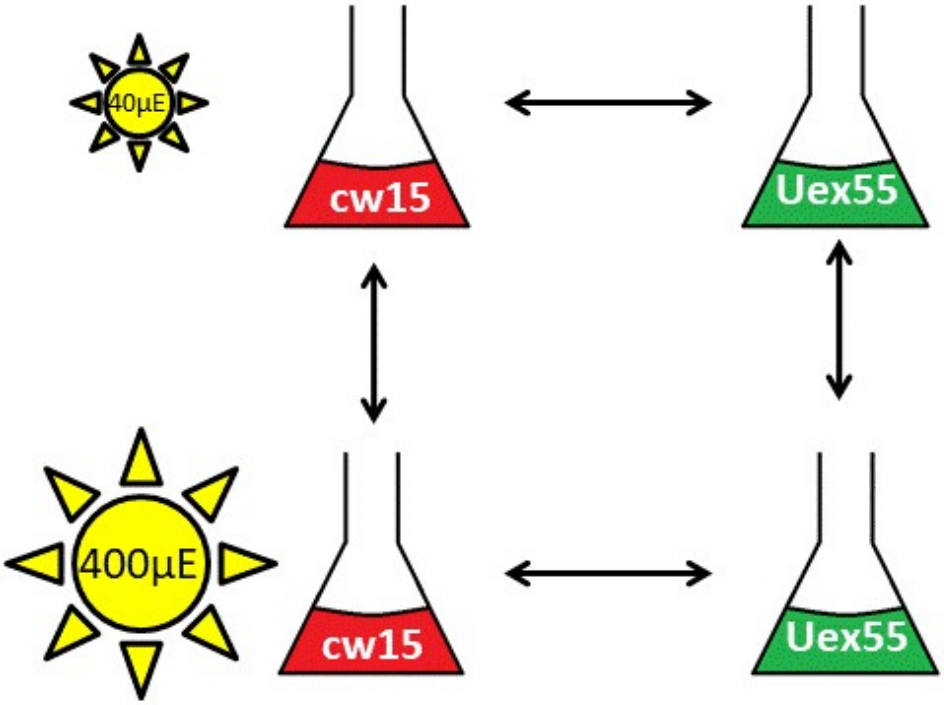


Figure 6.1: Diagram representing the 4 sample conditions and which comparisons were considered in the gene expression analysis of this study. For each C. reinhardtii strain (cw15 and Uex55) a comparison between SL (40 µmol $m^{-2}s^{-1}$) and HL (400 µmol $m^{-2}s^{-1}$) was performed, and at each light intensity a comparison between each strain was performed.

6.1.2. Aims of the study

- To analyse the change in transcriptome using RNA sequencing under a ten-fold increase in light intensity in *C. reinhardtii.*
- To determine if knocking out cytochrome c_{6A} alters the *C. reinhardtii* transcriptomic response to high light stress using RNA sequencing.

6.1.3. *C. reinhardtii* have unique mechanisms for motility and carbon concentration

It appeared there were possible effects on phototaxis and the carbon concentrating mechanism (CCM), so these topics will be reviewed briefly here before presenting the results.

6.1.3.1. Chlamydomonas reinhardtii exhibits both positive and negative phototaxis through flagella

Chlamydomonas reinhardtii is a freshwater alga, and therefore is exposed to varying light intensities due to weather or changing depth in the water column. *C. reinhardtii* is also unicellular, and uses two flagella to exhibit both negative and positive phototaxis to reach the optimal level of light intensity and respond to fluctuating light (Mayer 1968). When *C. reinhardtii* swims forwards it rotates, which allows periodic light exposure to a sensitive structure known as the eyespot organelle. Two channelrhodopsins (named chlamyopsins) located near the eyespot are thought to be responsible for detecting light intensity and in turn communicate with the motility machinery (Suzuki *et al.*, 2003; Berthold *et al.*, 2008). When the light intensity is too high or low, these chlamyopsins cause a depolarisation and action potential, which in turn causes the cell to change direction via Ca^{2+} signalling (Josef *et al.*, 2006; Arrieta *et al.*, 2017). This phototactic 'steering' allows for both positive and negative phototaxis.

Recent research has focussed on the regulation of this phototactic system. Calcium-induced phosphorylation of the chlamyopsin CHR1/COP3 has been observed and shown to be a key feedback mechanism in regulating phototactic sensitivity (Böhm *et al.*, 2019). Using acrylamide-pendant phosphate binding tag SDS-PAGE the study showed multiple phosphorylation sites on COP3, and suggested that different patterns may regulate positive and negative phototaxis. The redox state of the cytoplasm has also been shown to affect the phototactic response in *C. reinhardtii* (Wakabayashi *et al*., 2011). In this study the addition of the ROS hydrogen peroxide was shown to trigger positive phototaxis, and the ROS quencher 4-hydroxy-2, 2, 6, 6-tetramethyl piperidine 1-oxyl was shown to induce negative phototaxis. This suggests that a oxidised state of the cytoplasm leads to movement toward light, and a reduced state leads to movement away. Phototaxis has also been shown to be diminished upon exposure to the potent PSII inhibitor DCMU (Takahashi and King, 2003), suggesting that there is a link between photosynthesis and the phototactic response. For example, a change in light intensity would alter the rate of photosynthesis, and thus the redox state of reducing equivalents in the whole cell (Forti *et al.*, 2003), which has been shown to alter the flagellar beat frequency and therefore turning of the organism (Wakabayashi *et al.*, 2006).

6.1.3.2. The carbon concentration mechanism in C. reinhardtii allows efficient carbon fixation in low CO_2 conditions

Ribulose bisphosphate carboxylase oxygenase (rubisco) catalyses the fixation of inorganic carbon in the form of CO_2 to produce two molecules of 3-phosphoglycerate (3PG) from the 5 carbon molecule

ribulose bisphosphate. Rubisco can also catalyse a reaction between ribulose bisphosphate and molecular oxygen (O_2), thus producing one molecule of 3PG and one molecule of 2-phosphoglycerate (2PG). An energy-intensive process known as photorespiration must then convert the 2PG to 3PG, which decreases the efficiency of photosynthesis (Wingler *et al.*, 2000). The ratio of CO_2 to O_2 in the environment around rubisco can affect the relative rate of carboxylation and oxygenation, as both gases compete for the same active site (Jordan and Ogren, 1984; Andersson and Backlund, 2008). Therefore many photosynthetic organisms, including *C. reinhardtii*, have evolved CCMs in order to increase the relative rate of rubisco carboxylation by actively increasing the concentration of CO_2 around the enzyme (Raven *et al*, 2008). As *C. reinhardtii* is an aquatic alga, its CCM predominantly involves the regulation of bicarbonate pumping into the cell, followed by the conversion of the bicarbonate to CO_2 through the activity of carbonic anhydrases (Moroney *et al.*, 1985; Spalding, 2007). In *C. reinhardtii* rubisco is packaged into a structure in the chloroplast known as the pyrenoid (Kuchitsu *et al.*, 1988; Morita *et al.*, 2008), and a recent study has shown the importance of a scaffold protein known as EPYC1 in conjugating the rubisco inside the CCM organelle (Mackinder *et al.*, 2016). Therefore the CCM functions in increasing CO_2 concentration in the pyrenoid.

A major group of proteins essential to CCM in *C. reinhardtii* are inorganic carbon transporters. These proteins control influx of bicarbonate across the cell membrane and plastid membranes to ensure the carbon can reach the pyrenoid (figure 6.2). The proteins High Light Activated 3(HLA3), Low Carbon-Induced 1 (LCI1), Low Carbon-Induced A (LCIA), Carbonic Anhydrase 3 (CAH3) and Low Carbon-Induced B (LCIB) have been found to be key in two CO_2 influx pathways: one that functions at low CO_2 (0.03-0.50 %), and one that functions at very low CO_2 (under 0.03 %) environmental concentrations (Wang *et al.*, 2015). Under low CO_2 conditions, any CO_2 that diffuses into the chloroplast is thought to be converted into bicarbonate by LCIB before being transported into the thylakoid for conversion to CO_2 by the carbonic anhydrase CAH3. Under very low CO_2 conditions, HLA3 and LCIA are thought to assist in the transport of bicarbonate from outside the cell into the chloroplast, with HLA3 and LCIA being inorganic carbon transporters across the plasma and plastid outer membranes respectively. The identity of the thylakoid membrane inorganic carbon transporter has recently been suggested to be bestrophin 1 (BST1) (Mukherjee *et al.*, 2019). Another transporter found to be important in CCM is CEMA, a proton pump that actively moves protons from the stroma into the cytosol (Rolland *et al.*, 1997). This proton extrusion increases the pH of the stroma, allowing an optimal pH for carbonic anhydrase activity as well as increasing the concentration of bicarbonate relative to CO_2. This also suggests a strong link between compartmental pH and CCM.

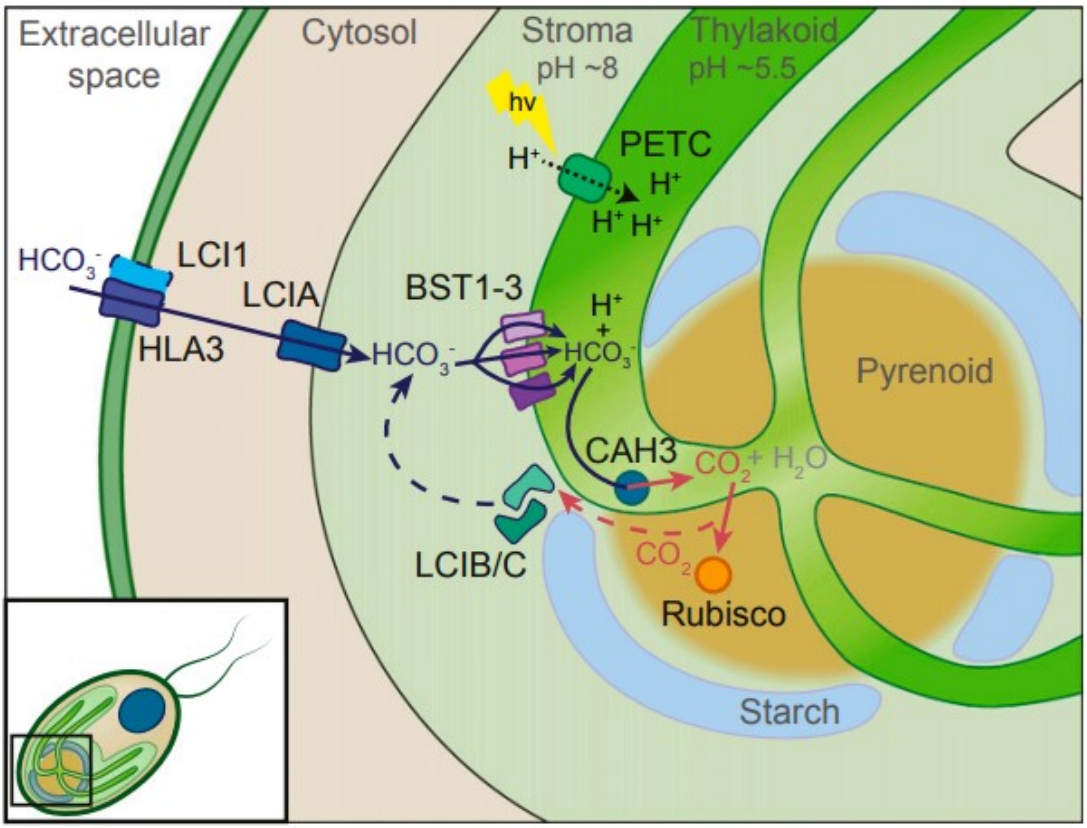


Figure 6.2: Overview of inorganic carbon transport across the plasma membrane, chloroplast envelope, thylakoid membrane and into the pyrenoid. Bicarbonate transported into the thylakoid lumen is converted into CO_2 in tubules running into the pyrenoid. The CO_2 then diffuses out of the thylakoid to concentrate around Rubisco. The photosynthetic electron transport chain (PETC) has been shown as a method of pumping protons into the lumen upon light exposure (hv). Taken from Mukherjee et al., 2019.

The expression of many CCM genes is under transcriptional control through retrograde signalling. There are many gaps in knowledge of the specifics of this signalling pathway. However, it is thought that Ca^{2+} may be involved, as a knockout mutant of the calcium-binding regulator CAS was shown to have diminished carbon concentrating capacity (Wang *et al.*, 2016). A transcriptional regulator known as CIA5 (CCM1) has been found to be responsible for upregulation of a majority of the CCM genes in response to low CO_2 levels as well as dark to light transitions (Xiang *et al.* 2001). Recent RNA sequencing data has identified the transcripts regulated by CIA5, which covered not only CCM machinery but around 25 % of all genes allocated in *C. reinhardtii* (Fang *et al.*, 2012). Another transcriptional regulator of CCM genes has been identified named LCR1, which has been shown to regulate CAH1, LCI1 and LCI6 (Yoshioka *et al.*, 2004). However the study showed that LCR1 has also

been shown to be dependent on CIA5, demonstrating that CIA5 appears to be the master regulator for CCM in *C. reinhardtii*. Interestingly, several mitochondrial genes have been shown to be upregulated under low CO_2 conditions (Fang *et al.*, 2012), therefore suggesting a role of the mitochondria in CCM too. This could be as a response to increased CCM as respiration does generate more CO_2, or could be linked to photorespiration regulation, as the mitochondria are heavily involved in the photorespiratory pathway. There are therefore gaps in the knowledge of CCM regulation that are yet to be resolved.

6.2. Results

6.2.1. Growth of C. reinhardtii cultures under high light stress and RNA extraction

Cultures of cw15 and the cytochrome c_{6A} knockout mutant line Uex55 were grown in triple biological replicates in TP medium under standard light conditions (40 µmol $m^{-2}s^{-1}$; SL) to mid-log growth phase. A sample was taken for RNA extraction before the remainder of the cultures were exposed to high light conditions (400 µmol $m^{-2}s^{-1}$; HL) for 4 hours. After this time another sample was taken from each culture. The RNA was then extracted and purified for RNA sequencing.

6.2.1.1. Quality control of extracted RNA

The quality of the extracted RNA was examined before any sample was sent for sequencing. First the concentration of RNA in each sample was estimated, along with the 260/230 and 260/280 UV absorbance ratios (figure 6.3). These give an idea of the purity of the sample, in particular any contamination from proteins or solvents such as phenol. All the RNA samples had very similar UV absorbance ratios, which for RNA should be around 2.0-2.3 for 260/230 and 2.0-2.1 for 260/280. This showed consistency and purity amongst the samples. All the RNA samples had over 150 ng/µL concentration; thus there was sufficient RNA for analysis.

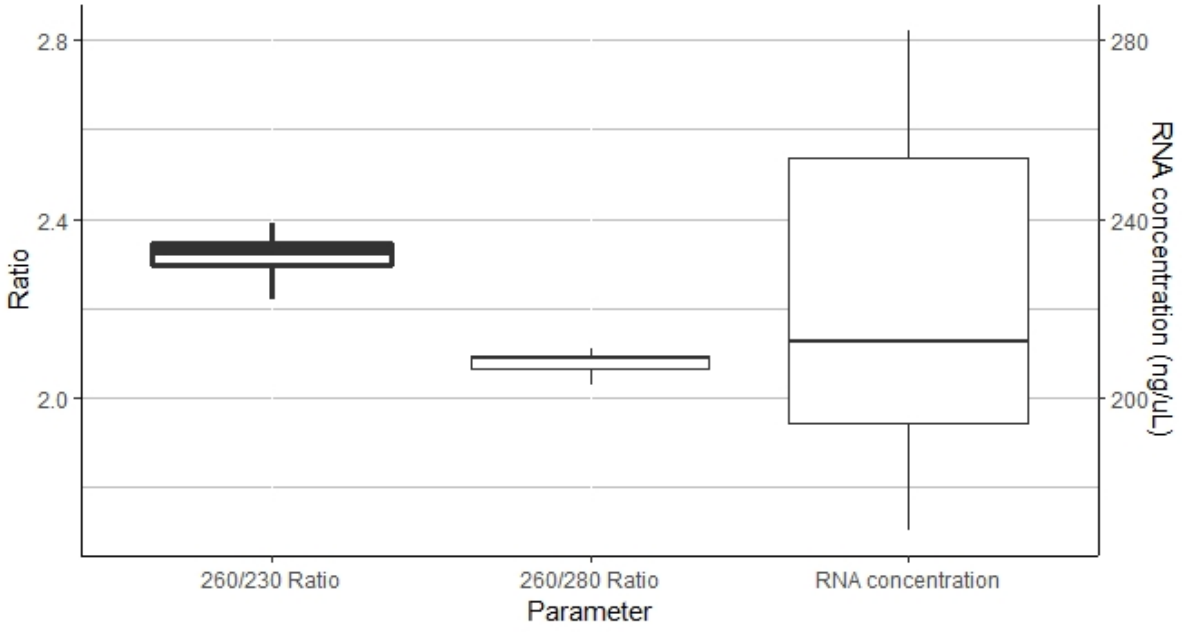


Figure 6.3: Boxplots of the 260/230 and 260/280 ratios, alongside the RNA concentrations measured, of the RNA samples taken. Each box represents the inter-quartile range, and the dividing line within the box represents the median value.

The RNA samples were also analysed by agarose gel electrophoresis to look for signs of degradation (figure 6.4). The 28S, 23S, 18S, 5S and tRNA bands were visible, showing that nuclear, mitochondrial and plastid rRNA was present with minimal sign of degradation. The 16S band was not strongly visible. The lanes looked consistent and therefore the samples were sent for RNA sequencing, performed at Novogene in Hong Kong.

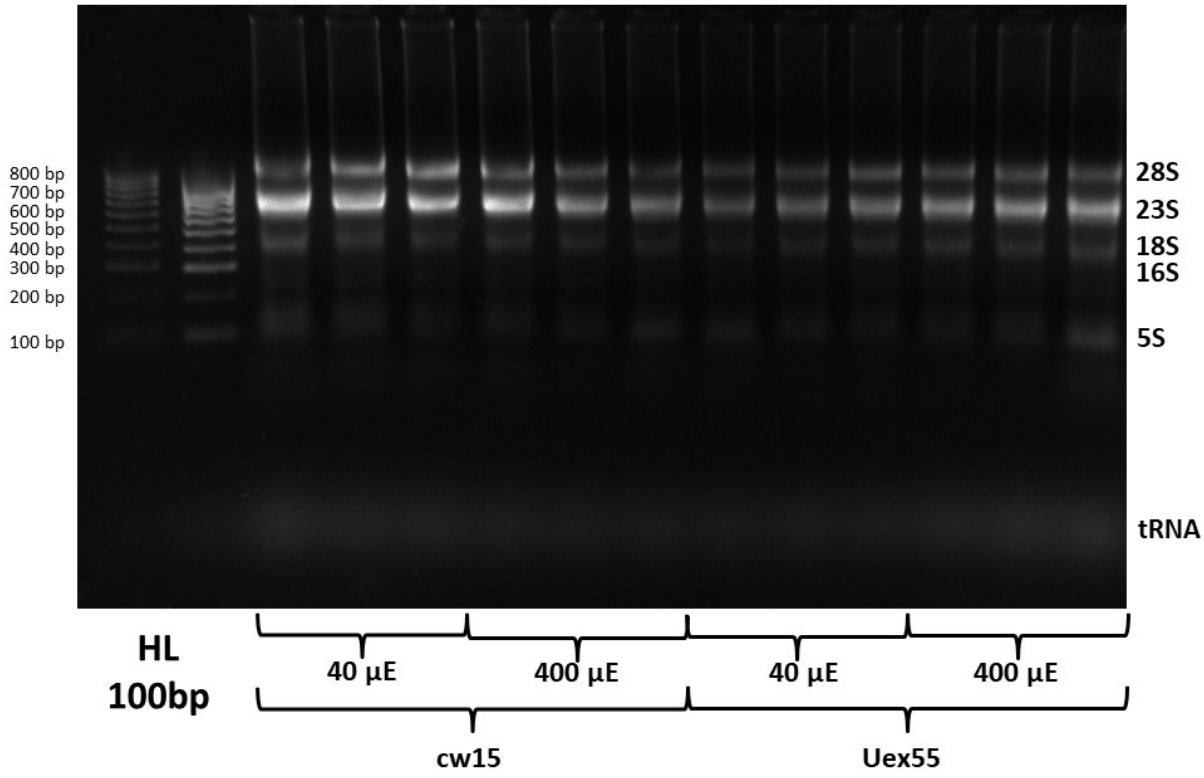


Figure 6.4: Electrophoresis gel of RNA samples. Hyperladder 100bp was used as a reference lane. The 28S, 23S, 18S, 16S, 5S and tRNA bands have been labelled.

6.2.1.2. Novogene quality control of extracted RNA and sequencing

The samples were quality control tested upon arrival at the RNA sequencing facility, to ensure that the samples passed their internal controls and to determine if significant degradation had occurred in transit from the UK to Hong Kong (supplementary figures S1 and S2, and Supplementary table S3). The RNA was analysed through nanodrop measurements and electrophoresis to ensure the purity and integrity of the RNA was acceptable for sequencing. The RNA concentration, and ratios of 260/280 and 260/230 had decreased slightly compared to the nanodrop analysis of figure 6.3, but not by enough to cause concern , especially when combined with the RNA integrity numbers (RIN), which were all above 8.5. The integrity of the electrophoresis bands did not appear to have changed

compared to the gel of figure 6.4., suggesting that minimal degradation occurred during sample transport. The 16S band is now visible, showing that 16S rRNA was present in the sample but at too lower concentration for detection in figure 6.4. Finally, the samples were tested using an Agilent 2100 bioanalyser and found to have high enough quality for RNA sequencing.

6.2.2. Quality control of RNA sequencing

Quality control of the sequencing process itself was also performed by Novogene. The quality of the reads was determined (supplementary figures S3-6 and supplementary table S4), with over 95 % of the reads in all samples having an error rate below 0.001 % (QC30), and an average error of around 0.02 % for all reads across each sample (QC20). The small percentage of reads that were not found to be 'clean' were reads containing adapters and thus not included in the final results. The GC content of the reads was also shown to be consistently around 63.5 %, which is very close to the calculated genomic GC content for *C. reinhardtii*, which is around 62 % (Sueoka; 1960). Overall this shows that both the RNA and RNA sequencing were of high enough quality for the results to qualify for analysis.

The reads were mapped to the JGI annotated *C. reinhardtii* genome v5.5 (Merchant *et al*., 2007) using Geneious, with around 79 % reads successfully mapped across each sample (data not shown). Transcript levels for annotated genes were calculated for each sample, and from this data relative transcript levels were calculated between specific sample sets: between cw15 and Uex55 at SL and HL, and between SL and HL for cw15 and Uex55 (figure 6.1).

6.2.3. Key findings of RNA sequencing

The first gene locus investigated in the RNA sequencing data was cytochrome c_{6A} itself (*CYC4*, cre16.g670950). No full transcripts were found mapped to cytochrome c_{6A} in any of the Uex55 samples, which was to be expected. In the transition from SL to HL in cw15, the $\log_2$ fold change in cytochrome c_{6A} transcripts was 0.687 on average, which is only a marginal increase, despite being statistically significant ($0=1.14 \times 10^{-4}$). As cytochrome c_{6A} has a low native transcript level (Nimmo, 2011), this change does not represent a large increase in the absolute number of cytochrome c_{6A} mRNA molecules. Therefore there may be a mild transcriptional regulation of cytochrome c_{6A} under high light stress, but the effect is likely to be small.

6.2.3.1. The top twenty largest changes in transcript level for each condition

For each comparison of the 4 samples shown in figure 6.1 (HL against SL in cw15 and Uex55, and cw15 against Uex55 in HL and SL) the 20 largest changes in transcript level (10 most downregulated and 10 most upregulated) were recorded. Any annotation, potential homologue or domain for each gene was then investigated to attempt to understand its role in *C. reinhardtii*.

6.2.3.1.1 Comparison between high light and standard light in cw15

The transcripts showing the largest changes in levels between SL to HL in cw15 are shown in table 6.1. Notable genes observed in this table as upregulated include many genes involved in carbon concentrating for *C. reinhardtii*, previously identified as being upregulated in low CO_2 conditions: *CCP1, CCP2, LCI23, CAH4, CAH5, LCIA* and *LCI1*. As the cultures were grown without bubbling with extra air, the conditions could be seen as CO_2-limited as well as under high light stress. There may also be a regulation of growth occurring, with cell wall degradation enzyme autolysin being downregulated. The downregulated epidermal growth factor (EGF)-like protein may also be involved in cell proliferation. This may suggest an increase in growth rate under high light. It was also noted that a flagellar associated protein (FAP) was downregulated upon high light exposure, suggesting a modulation of motility and therefore phototaxis. Aurora-like kinases have also been implicated in flagellar assembly and disassembly in *C. reinhardtii* (Pan *et al.*, 2004; Luo *et al.*, 2011), further suggesting modulation of motility.

Locus/ID	Log_2 differential transcript levels under high light	P-value	Notes
Cre01.g042100	-4.86	1.44E-77	Non-specific serine/threonine kinase
Cre07.g355950	-4.23	5.53E-38	DNAJ protein
Cre07.g354750	-4.11	2.15E-27	Hypothetical protein
Cre13.g603550	-3.98	6.69E-28	Hypothetical protein
Cre09.g400590	-3.83	4.33E-113	EF-hand protein
Cre16.g668700	-3.75	6.65E-80	EGF-like protein
Cre07.g326700	-3.74	8.62E-78	Potential transcription factor
Cre14.g623125	-3.62	6.17E-61	Flagellar associated protein
Cre14.g616700	-3.60	2.95E-18	Aromatic amino acid transaminase
Cre01.g025200	-3.54	7.33E-44	Autolyisn/Gametolysin

Cre04.g223300	5.016	1.74E-25	*CCP1*
Cre02.g112366	5.47	5.84E-31	Hypothetical protein
Cre04.g222750	5.52	2.19E-38	*CCP2*
Cre02.g112333	5.59	2.96E-31	Aurora kinase
Cre06.g281600	5.62	2.94E-29	*LCI23*
Cre05.g248400	5.75	6.73E-23	*CAH4* carbonic anhydrase
Cre05.g248450	5.88	2.06E-24	*CAH5* carbonic anhydrase
Cre06.g309000	6.67	1.11E-55	*LCIA* anion transporter
Cre12.g555700	7.26	1.64E-45	*DNJ15* DNAJ protein
Cre03.g162800	7.63	4.16E-47	*LCI1*

Table 6.1: Largest log_2 changes in transcript level upon high light exposure for 4 h in cw15. Gene loci shown in red represent a lower transcript level upon HL exposure, and gene loci shown in blue represent a high transcript level upon HL exposure. Potential role or gene name is given in "notes".

6.2.3.1.2. Comparison between high light and standard light in Uex55

The transcripts showing largest changes in levels between HL and SL in Uex55 are shown in table 6.2. Interestingly, there is only one gene that can be found in both table 6.1 and 6.2, Cre07.g354750, encoding a hypothetical protein. This gene showed downregulation at the transcript level upon HL exposure in both cw15 and the knockout line. Unfortunately no annotations were found for this locus, and homology modelling and BLAST yielded no significant results. However, the change in level of this transcript does suggest that some genes were regulated in a similar way under high light stress in both cw15 and Uex55. Many affected genes in both Uex55 and cw15 under HL stress appear to be involved in transcriptional control, with transcript levels for zinc finger proteins, histone acetyltransferases and histone methyltransferases being altered. Translational and post-translational control were also affected, with transcripts for kinases, ring finger domain proteins and elongation factors demonstrating regulation under high light stress. Without more information on the activators or targets of these gene products it is difficult to determine precisely which system these genes are involved in. However, it should be noted that regulation under HL stress did still occur in the Uex55 knockout line. The downregulation of a chlorophyll *a-b* binding protein may suggest a downregulation of chlorophyll biosynthesis, and an upregulation of *VTE3* could result in an increase in plastoquinone as the MPBQ/MSBQ methyltransferase it encodes is involved in PQ synthesis (Cheng *et al*., 2003).

Locus/ID	Log_2 differential transcript levels under high light	P-value	Notes
Cre01.g009101	-3.55	6.25E-42	Ring finger domain protein
Cre03.g212977	-2.79	1.31E-34	Histone acetyltransferase
Cre09.g394325	-2.75	2.28E-51	Chlorophyll a-b binding protein
Cre07.g354750	-2.38	7.62E-16	Hypothetical protein
Cre06.g268850	-2.27	2.17E-21	Serine threonine protein kinase
Cre03.g163950	-2.22	4.05E-28	Cysteine dioxygenase
Cre02.g086000	-2.21	7.52E-16	*ARS8* Arylsulfatase
Cre17.g726526	-2.20	2.05E-16	Ferrochelatase homologue
Cre11.g479900	-2.19	6.85E-33	Thioredoxin homologue
Cre06.g274050	-2.19	4.28E-43	Histone lysine N-methyltransferase
Cre07.g315050	2.07	2.58E-21	*GGH1* Gamma-glutamyl hydrolase
Cre14.g625450	2.09	5.90E-33	*VTE3* MPBQ/MSBQ methyltransferase
Cre17.g719150	2.11	2.48E-08	Zinc finger protein
Cre13.g586000	2.16	2.32E-33	*ACSF3* malonyl-CoA/methylmalonyl-CoA synthetase
Cre17.g698850	2.16	1.47E-33	Starch-debranching enzyme
Cre10.g446350	2.16	6.65E-18	PSbP domain protein
Cre06.g306900	2.21	1.91E-10	*SRR9* Scavenger receptor cysteine rich protein
Cre12.g559800	2.23	2.23E-34	Glutathione S transferase
Cre10.g435500	2.33	2.21E-19	*EFT2* Elongation factor Ts-like protein
Cre14.g626300	2.63	1.92E-16	Amidase homologue

Table 6.2: Largest log_2 changes in transcript level upon high light exposure for 4 h in Uex55. Gene loci shown in red represent a lower transcript level upon HL exposure, and gene loci shown in blue represent a high transcript level upon HL exposure. Potential role or gene name is given in "notes".

6.2.3.1.3. Comparison between Uex55 and cw15 under high light conditions

The transcripts showing largest differences in levels between Uex55 and cw15 under HL are shown in table 6.3. Many kinases, including MAPKKK and CDK-like proteins appeared to have higher transcript levels in Uex55 compared to cw15. One such kinase appears to be involved in flagellin

binding, suggesting again a regulation of motility in response to high light to different extents in cw15 against Uex55. Two closely related nitrate transporter genes (*NRT2.4* and *NRT2.5*) also had a higher transcript level in Uex55 than cw15. Transcriptional control genes including a histone acetyltransferase and a zinc finger protein were shown to have higher transcript levels in cw15 under HL exposure, as well as protein-protein interaction factors such as Ankyrin repeat, DNAJ and tetratricopeptide repeat proteins and the heat shock transcription factor *HSF2*. A gene involved in cross-membrane transport of amino acids (*BAT1*) also had higher transcript levels in cw15 than Uex55 under HL conditions, as well as bestrophin 1 (*BST1*), a gene recently suggested to encode a thylakoid membrane inorganic carbon transporter (Mukherjee *et al.*, 2019). It is worth noting that the differential transcript level value does not specify whether, for example, a high log differential transcript level value of Uex55 against cw15 would be due to a high transcript level in the former, a low transcript level in the latter, or both. These data suggest that there was a difference in regulation of transcription under HL stress depending on whether or not cytochrome c_{6A} was present.

Locus/ID	Log_2 differential transcript level of Uex55 relative to cw15	P-value	Notes
Cre07.g331700	-3.64	1.26E-98	Hypothetical protein
Cre01.g023650	-3.63	3.56E-20	BAT1 amino acid permease
Cre03.g212977	-3.59	4.90E-21	Histone acetyltransferase
Cre06.g278129	-3.46	7.16E-40	CGI-12 protein
Cre10.g431900	-3.44	4.18E-50	Ankyrin repeat protein
Cre16.g662600	-3.22	2.04E-47	Bestrophin (chloride channel)
Cre05.g246650	-3.00	1.84E-28	SF4 Zinc finger protein
Cre07.g354500	-2.97	3.30E-125	HSF2 Heat shock transcription factor
Cre09.g389615	-2.88	2.11E-63	Tetratricopeptide repeat protein
Cre09.g395732	-2.82	1.82E-40	DNAJ protein
Cre03.g150151	3.13	5.75E-17	NRT2.5 Nitrate transporter
Cre03.g150101	3.25	1.50E-19	NRT2.4 Nitrate transporter
Cre10.g432150	3.46	2.95E-58	Cyclin dependent kinase-like protein
Cre09.g410332	3.69	7.30E-81	Serine-threonine protein kinase
Cre03.g173100	3.71	8.99E-40	Hypothetical kinase
Cre10.g435750	3.76	3.50E-82	Hypothetical protein
Cre02.g141826	3.79	1.13E-35	MAPKKK
Cre07.g355950	3.92	6.51E-34	DNAJ protein
Cre01.g042100	3.94	3.20E-44	Possible flagellin-binding non-specific protein kinase
Cre09.g399886	4.79	4.84E-77	Reverse transcriptase

Table 6.3: Largest log_2 differences in transcript level of Uex55 relative to cw15 under HL conditions. Gene loci shown in red represent a lower transcript level in Uex55 relative to cw15 upon HL exposure, and gene loci shown in blue represent a lower transcript level in cw15 relative to Uex55 upon HL exposure. Potential role or gene name is given in "notes".

6.2.3.1.4. Comparison between Uex55 and cw15 under standard light conditions

The transcripts with the largest differences in levels between Uex55 and cw15 under SL are shown in table 6.4. As with the HL condition, many of the transcripts that were more abundant in Uex55 than

cw15 encoded kinases. Two kinases were also shown to have higher transcript levels in cw15 than Uex55. Additionally, ankyrin repeat protein and zinc finger protein both had higher transcript levels in cw15. These observations suggest that again there was a difference in gene expression depending on whether cytochrome c_{6A} was present, but these were not necessarily the same as those under HL stress. The potential protease TRAB family protein was also found to have higher transcript levels in cw15 under SL. In prokaryotes TRAB proteins were found to play a role in intercellular communication and information transfer (An and Clewell; 1994), and therefore are thought to also play a signalling role in plants (Duncan *et al.*, 2013). Various metabolic pathways appeared to be affected differently in cw15 and Uex55, with sucrose 1-fructosyltransferase and 2-oxoglutarate-dependent dioxygenase having higher transcript levels in cw15, and alcohol dehydrogenase and 8-hydroxy-5-deazaflavin:NADPH oxidoreductase having higher transcript levels in Uex55. Therefore differences in Uex55 and cw15 metabolic control, even in the absence of high light stress, were inferred.

Locus/ID	Log_2 differential transcript levels of Uex55 relative to cw15	P-value	Notes
Cre10.g431900	-3.16	3.64E-129	Ankyrin repeat protein
Cre05.g246650	-2.61	1.36E-96	SF4 Zinc finger protein
Cre09.g412350	-2.23	1.56E-107	2-oxoglutarate-dependent dioxygenase
Cre10.g463370	-2.16	2.57E-87	Myosin light chain kinase
Cre02.g104450	-1.65	6.60E-27	Tyrosine kinase
Cre01.g016700	-1.34	5.00E-40	Microtubule-associated protein
Cre02.g095121	-1.31	2.29E-33	16S rRNA pseudouridine synthase
Cre12.g488000	-1.31	6.63E-39	Sucrose 1-fructosyltransferase
Cre14.g633500	-1.30	1.70E-41	Contains DENN domain 6A
Cre16.g677800	-1.29	1.52E-14	TRAB family protein
Cre01.g038350	1.17	4.37E-13	Hypothetical early growth phase protein
Cre03.g206200	1.22	3.94E-32	Hypothetical protein
Cre14.g623650	1.32	5.93E-24	Alcohol dehydrogenase ($NADP^+$)
Cre01.g039702	1.49	1.67E-88	Crinkler family protein, potential kinase
Cre02.g141886	1.51	2.76E-18	MEKK Tyrosine kinase
Cre14.g631600	1.52	1.18E-29	8-hydroxy-5-deazaflavin:NADPH oxidoreductase
Cre16.g687400	1.97	1.38E-106	Cyclin dependent kinase-like protein
Cre02.g141826	2.27	4.47E-43	MAPKKK
Cre03.g179350	2.60	4.56E-142	MFS1 Major facilitator superfamily
Cre09.g399886	3.37	1.32E-112	Reverse transcriptase

Table 6.4: Largest log_2 changes in transcript level of Uex55 relative to cw15 under SL conditions. Gene loci shown in red represent a lower transcript level in Uex55 relative to cw15 under SL, and gene loci shown in blue represent a lower transcript level in cw15 relative to Uex55 under SL. Potential role or gene name is given in "notes".

6.2.3.2. Calculating the largest difference in transcript regulation between Uex55 and cw15 upon high light exposure implicates regulation of carbon concentrating mechanisms and motility.

In order to determine which genes had the largest difference in transcript level changes upon HL exposure between cw15 and Uex55, the $\log_2$ fold change of the latter on transfer to HL was subtracted from that of the former. The transcripts showing the largest differences in either direction are summarised in table 6.5. For a majority of the large differences in HL regulation between cw15 and Uex55, it appeared that the differences observed were due to regulation occurring in the wild type but not in Uex55. This suggests that induced changes under high light were negated or suppressed without cytochrome c_{6A} present, implying a signalling role.

Locus/ID	cw15 High light vs low light	Uex55 High light vs low light	Calculated difference	Notes
Cre04.g223300	5.02	0.90	4.12	*CCP1* metabolite transporter
Cre04.g222750	5.52	0.85	4.67	*CCP2* carbon concentration
Cre06.g281600	5.62	0.87	4.75	*LCI23* carbon concentration
Cre02.g112366	5.47	0.64	4.83	Hypothetical protein
Cre02.g112333	5.59	0.61	4.97	Aurora kinase
Cre05.g248400	5.75	0.57	5.18	*MCA/mtCA1* carbonic anhydrase
Cre05.g248450	5.88	0.60	5.27	*CAH5* Carbonic anhydrase
Cre06.g309000	6.67	0.71	5.96	*LCIA/NAR1.2* Formate/nitrate transporter
Cre12.g555700	7.26	0.45	6.81	*DNJ15* DNAJ protein
Cre03.g162800	7.63	0.32	7.31	*LCI1* carbon concentration
Cre01.g042100	-4.86	0.27	-5.12	Serine/threonine kinase
Cre10.g435750	-3.16	1.42	-4.59	MAPKKK
Cre09.g410332	-3.30	1.25	-4.56	Serine/threonine kinase
Cre07.g355950	-4.23	0.19	-4.42	DNAJ protein
Cre14.g623125	-3.62	0.34	-3.96	Flagellar associated protein
Cre03.g173100	-3.48	0.44	-3.92	Laminarinase
Cre17.g736511	-2.64	1.23	-3.87	Hypothetical protein
Cre10.g432150	-3.16	0.67	-3.83	CDK-like protein
Cre03.g150151	-2.97	0.82	-3.79	*NRT2.5* Nitrate transporter
Cre03.g150101	-3.44	0.33	-3.78	*NRT2.4* Nitrate transporter

Table 6.5: Twenty largest calculated differences between the transcriptional change under HL exposure of cw15 and Uex55. The difference was calculated by subtracting the log_2 change from SL to HL for Uex55 from that of cw15. Loci coloured red were genes that showed a larger increase in transcript level in cw15 than Uex55 on transfer to HL, and loci coloured blue were genes that showed a larger decrease in transcript level in cw15. Potential role or gene name is given in "notes".

Many of the genes with lower changes in transcript levels in cw15 than Uex55 were for kinases involved in post-translational modification. The specific proteins that these kinases target are unknown. However, this does imply that protein activity was being modulated at a post-translational level in addition to transcriptional. A flagellar associated protein (FAP) was also found, again suggesting a downregulation in motility only seen when cytochrome c_{6A} was present. Therefore the transcript levels of genes given the Mapman ontology of Eukaryotic motility were investigated (figure 6.5).

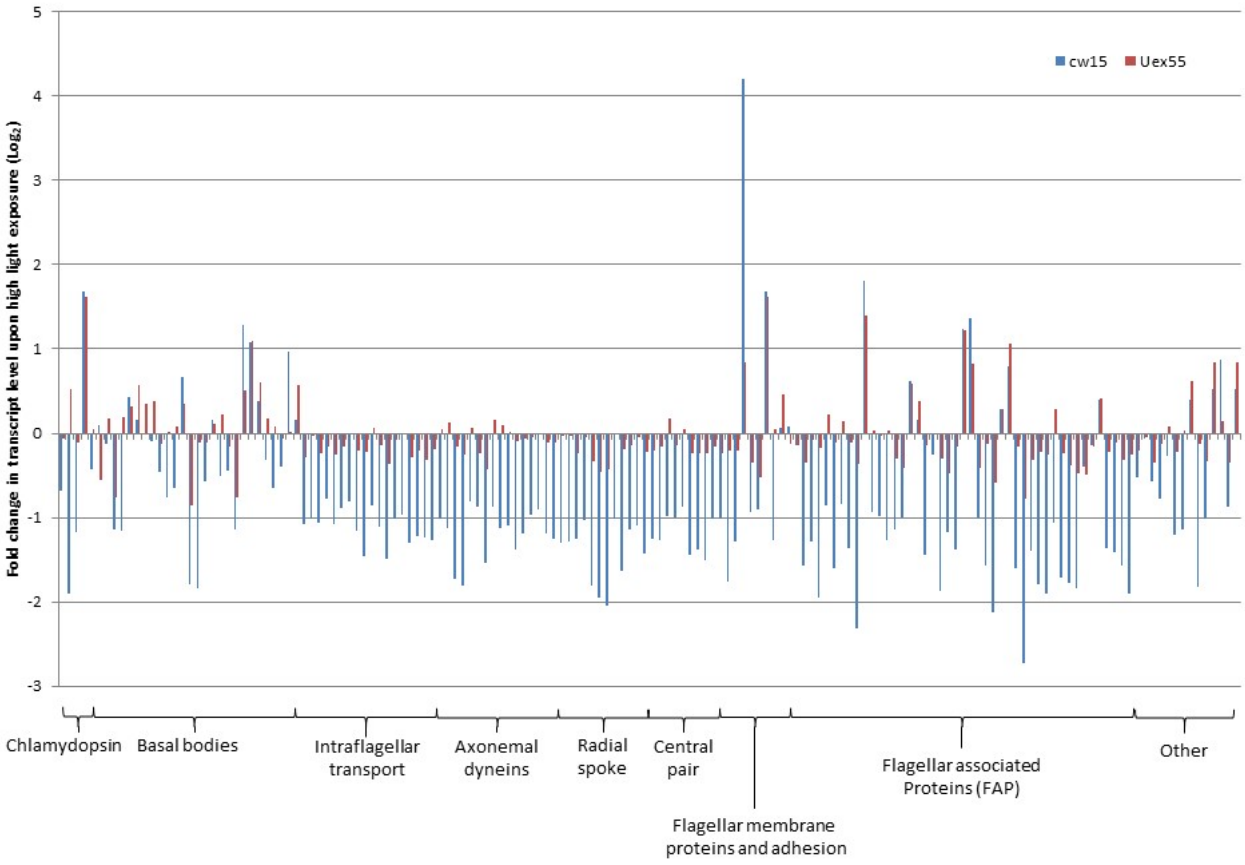


Figure 6.5: Comparison of transcript level change upon high light exposure for 4 h in cw15 (blue) and Uex55 (red) for all transcripts with the Mapman ontology "Eukaryotic motility". The genes are ordered by more specific Mapman clusters, which are denoted at the bottom of the graph.

There appeared to be an overall trend of downregulation in cw15 of motility proteins in response to HL, which was suppressed in Uex55. This is especially seen in the transcripts related to intraflagellar transport, axonemal dyneins, radial spoke, central pair and FAP. Intraflagellar transport genes function in the transport of subunits along the axoneme, and therefore are required for flagellar construction (Cole, 2003). A putative MAPK named *LF4* (cre13.g582650) has been identified as essential for flagellum length in *C. reinhardtii* (Berman *et al.*, 2003), and this too was downregulated in cw15 under HL exposure but not in Uex55 (data not shown). Therefore, these data imply an overall downregulation of flagellar subunit expression and construction under HL conditions in *C. reinhardtii*, which was reliant either directly or indirectly on cytochrome c_{6A} being present.

The genes that had higher elevation of transcript levels upon HL exposure in cw15 compared to Uex55 showed an enrichment of proteins involved in carbon concentrating mechanisms, including carbonic anhydrases. Therefore the transcript levels of genes with the Mapman ontology "carbon concentrating mechanism algal" were investigated under HL exposure (figure 6.6A). Within this ontology the increase in transcript levels of *LCID*, *CCP1*, *CCP2*, *LCI23*, *LCI1*, *LCIE* and *CEM1* under HL stress was diminished in Uex55. *CCP1* and *CCP2* are thought to encode proteins needed for inorganic carbon transport across the chloroplast envelope (Pollock *et al.*, 2004), and *LCI1* has been suggested to encode a bicarbonate intake protein at the cell membrane (Onishi *et al.*,2010). *CEM1* (also named *CEMA*) was found to encode a proton pump that moves hydrogen ions from the stroma to the cytosol to maintain a high stromal pH and allow further bicarbonate import (Rolland *et al.*, 1997). Relatively little is known about *LCID*, *LCI23* and *LCIE* except that they are induced under low CO_2 conditions and are linked to inorganic carbon acquisition (Wang *et al*., 2015). There may also be a slightly diminished upregulation of the transcript for the large subunit of rubisco (RbcL), but this was found to be less pronounced than the differences for inorganic carbon transport transcripts.

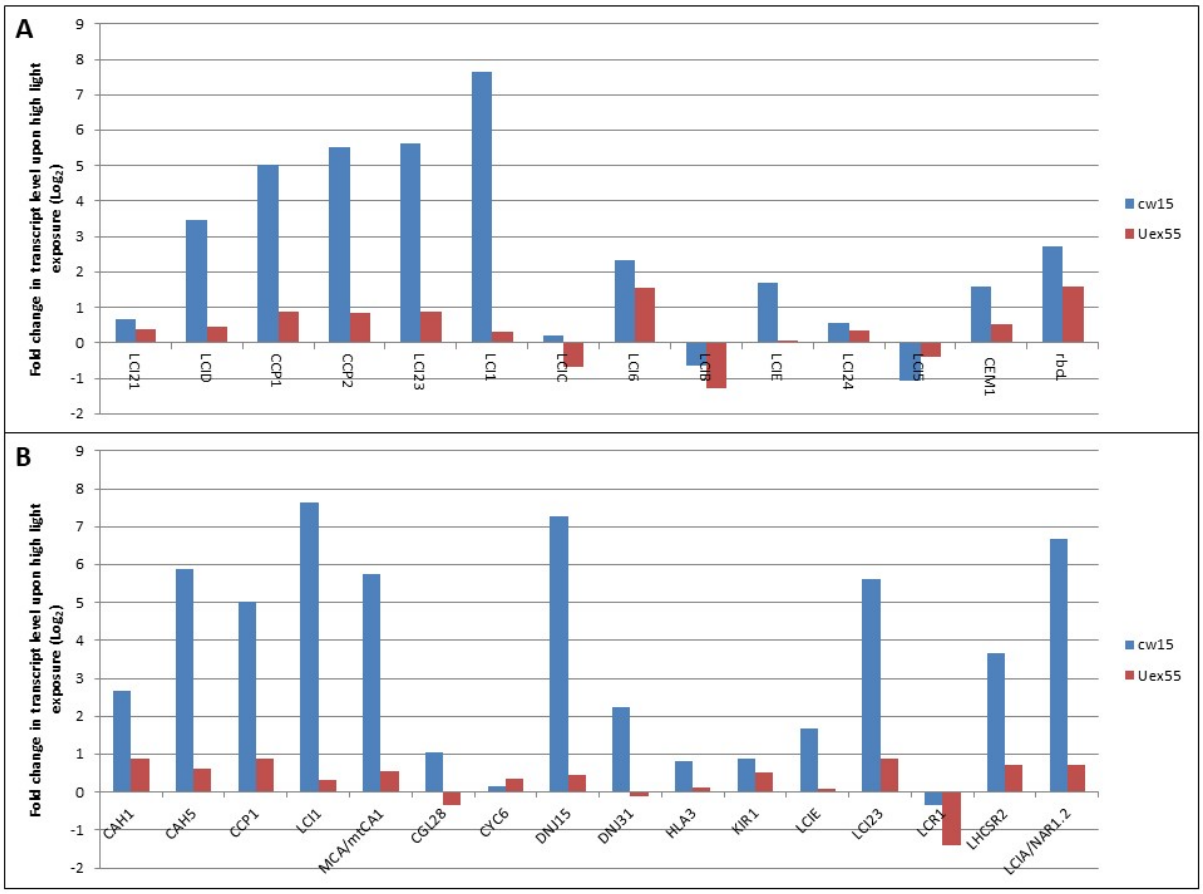


Figure 6.6: Comparison of transcript level change upon high light exposure for 4 h in cw15 (blue) and Uex55 (red) for all transcripts with the Mapman ontology "Carbon concentration mechanism algal" (A), and for transcripts of genes identified in the carbon concentration mechanism ontology cluster controlled by CCM1 as identified in Fang et al., 2012 (B).

The carbon concentration mechanism genes in *C. reinhardtii* are predominantly under the control of the master switch CIA5 (also named CMM1; Miura *et al*., 2004). RNA sequencing of a *C. reinhardtii* mutant strain with *CIA5* knocked out through a point mutation separated the genes controlled by *CIA5* into several clusters by ontology (Fang *et al.*, 2012). One cluster was named as "carbon concentrating mechanism" and was noted for containing a high proportion of transporters. Transcript levels of genes in this cluster were therefore analysed in this study (figure 6.6). Around 81 % of the genes in this cluster were also found to have diminished regulation in the absence of cytochrome c_{6A}. This could suggest a signalling link between *CIA5* and cytochrome c_{6A}, although it must be noted that the former controls transcription of a wider range of genes than just CCM, and the range does not perfectly overlap with that of cytochrome c_{6A}.

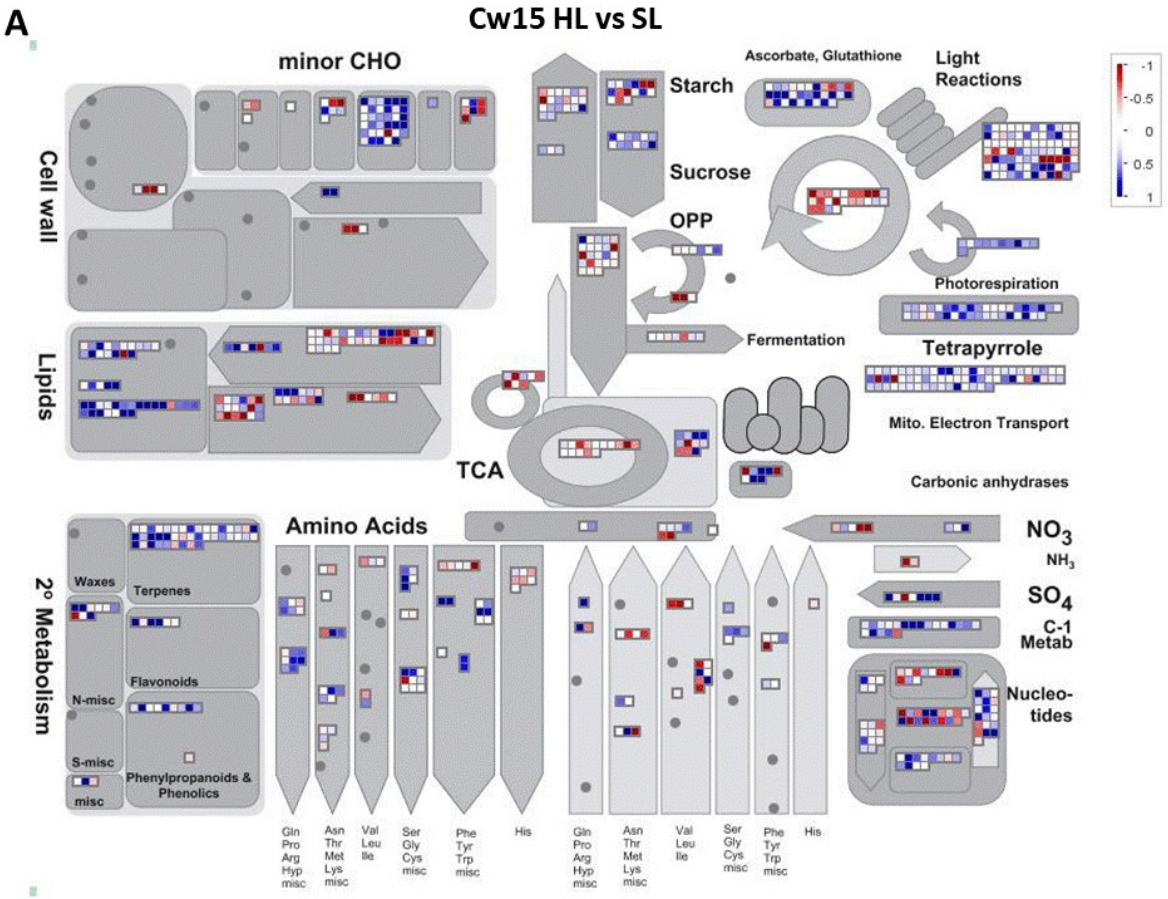
Therefore these data suggest that cytochrome c_{6A} plays a role in control of nuclear transcript levels in response to high light stress, in particular a downregulation of flagellar motility and upregulation in carbon concentration on transfer to high light.

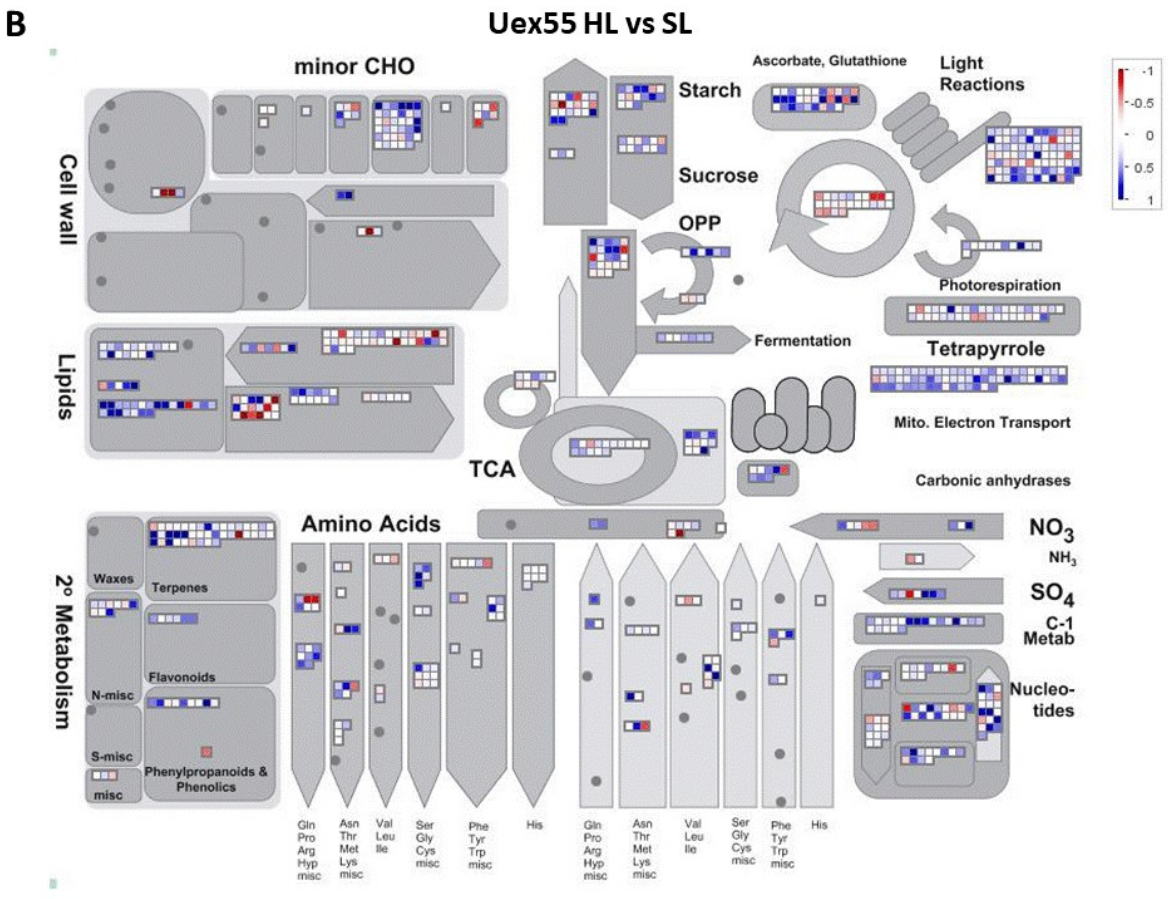
6.2.3.3. Analysing the RNA sequencing data using Mapman

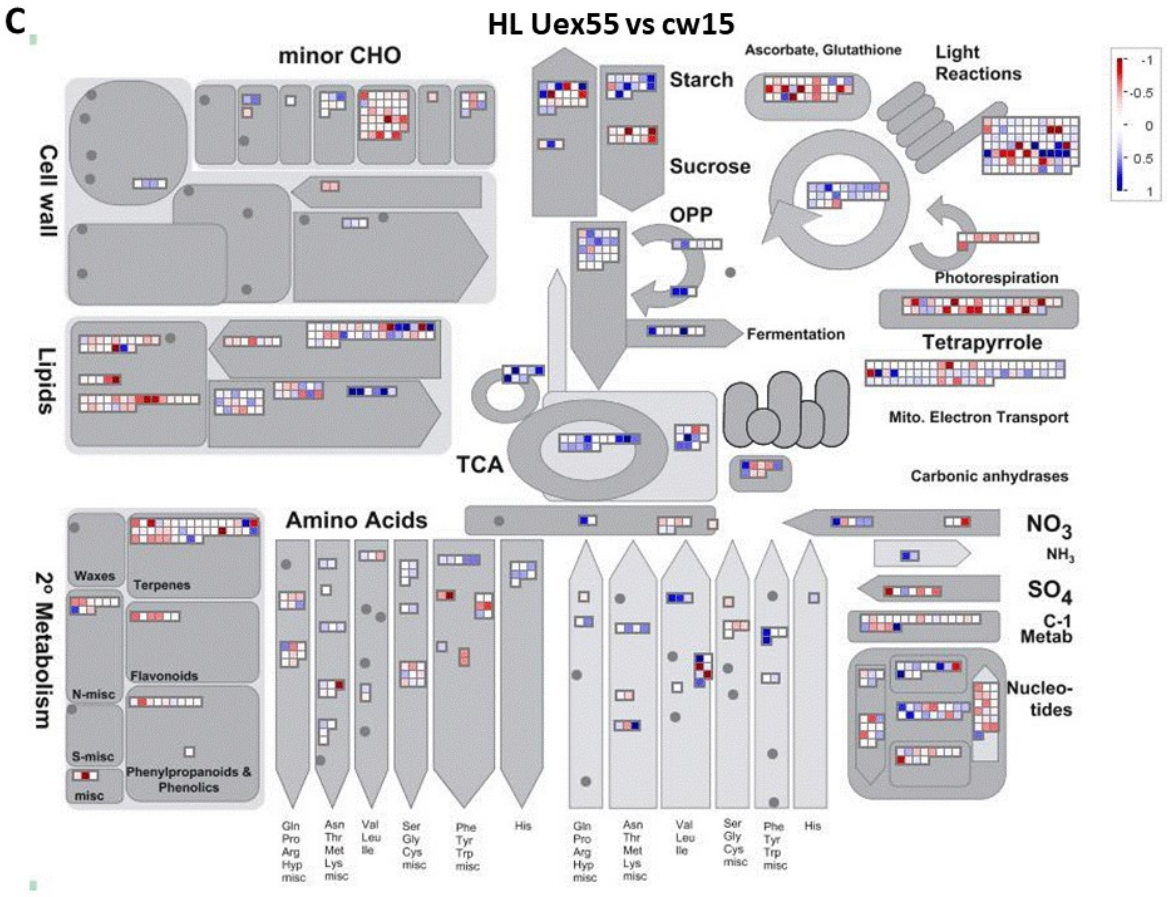
To get a broader view of the effects of cytochrome c_{6A} on the transcriptional expression of specific pathways and functions of *C. reinhardtii* under HL stress, the RNA sequencing data were also analysed using Mapman software (Thimm *et al.*, 2004).

6.2.3.3.1. Metabolic overview

The effect of knocking out cytochrome c_{6A} on the *C. reinhardtii* response to HL stress on metabolism as a whole was analysed (figure 6.7). Under HL conditions in cw15 (figure 6.7A) there is a general upregulation of many synthetic pathways, including tetrapyrrole, cell wall precursor, phospholipid and amino acid synthesis. As these pathways would use resources such as ATP and reducing equivalents from photosynthesis, upregulating these pathways would provide a sink for energetic metabolites to help drive photosynthesis and increase photochemical quenching. This upregulation of anabolic pathways is also seen in Uex55 upon HL exposure (figure 6.7B), suggesting that cytochrome c_{6A} does not play a role in regulating these pathways directly or indirectly. This complements the chlorophyll fluorescence data of this study that found no significant difference in photochemical quenching when cytochrome c_{6A} was knocked out.


A
Cw15 HL vs SL
minor CHO
Cell wall
Lipids
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
misc
Phenylpropanoids & Phenolics
Starch
Sucrose
OPP
Fermentation
TCA
Ascorbate, Glutathione
Light Reactions
Photorespiration
Tetrapyrrole
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
-1
-0.5
0
0.5
1


B
Uex55 HL vs SL
minor CHO
Cell wall
Lipids
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
misc
Phenylpropanoids & Phenolics
Starch
Sucrose
OPP
Fermentation
TCA
Ascorbate, Glutathione
Light Reactions
Photorespiration
Tetrapyrrole
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
-1
-0.5
0
0.5
1


C
HL Uex55 vs cw15
minor CHO
Cell wall
Lipids
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
misc
Phenylpropanoids & Phenolics
Starch
Sucrose
OPP
Fermentation
TCA
Ascorbate, Glutathione
Light Reactions
Photorespiration
Tetrapyrrole
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
-1
-0.5
0
0.5
1

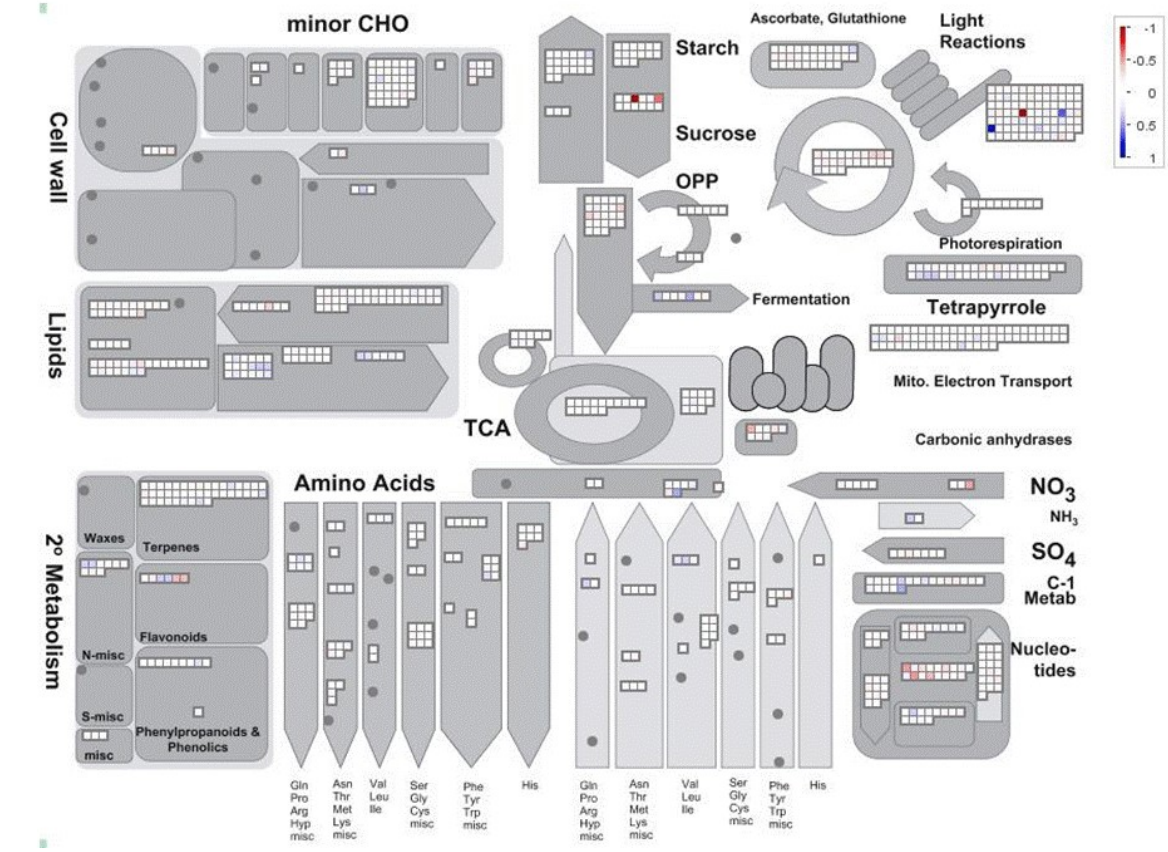

D
SL Uex55 vs cw15
minor CHO
Cell wall
Lipids
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
misc
Phenylpropanoids & Phenolics
Starch
Sucrose
OPP
Fermentation
TCA
Ascorbate, Glutathione
Light Reactions
Photorespiration
Tetrapyrrole
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
-1
-0.5
0
0.5
1

Figure 6.7: A comparison of transcript levels of cw15 and Uex55 under standard (40 µmol.s^{-1}.m^{-2}) and high (400 µmol.s^{-1}.m^{-2}) light conditions in the C. reinhardtii Mapman pathway "Metabolic overview". The blue squares represent a higher transcript level under high light stress relative to standard light (A and B) or higher transcript levels in Uex55 compared to cw15 (C and D), whereas red squares represent the opposite.

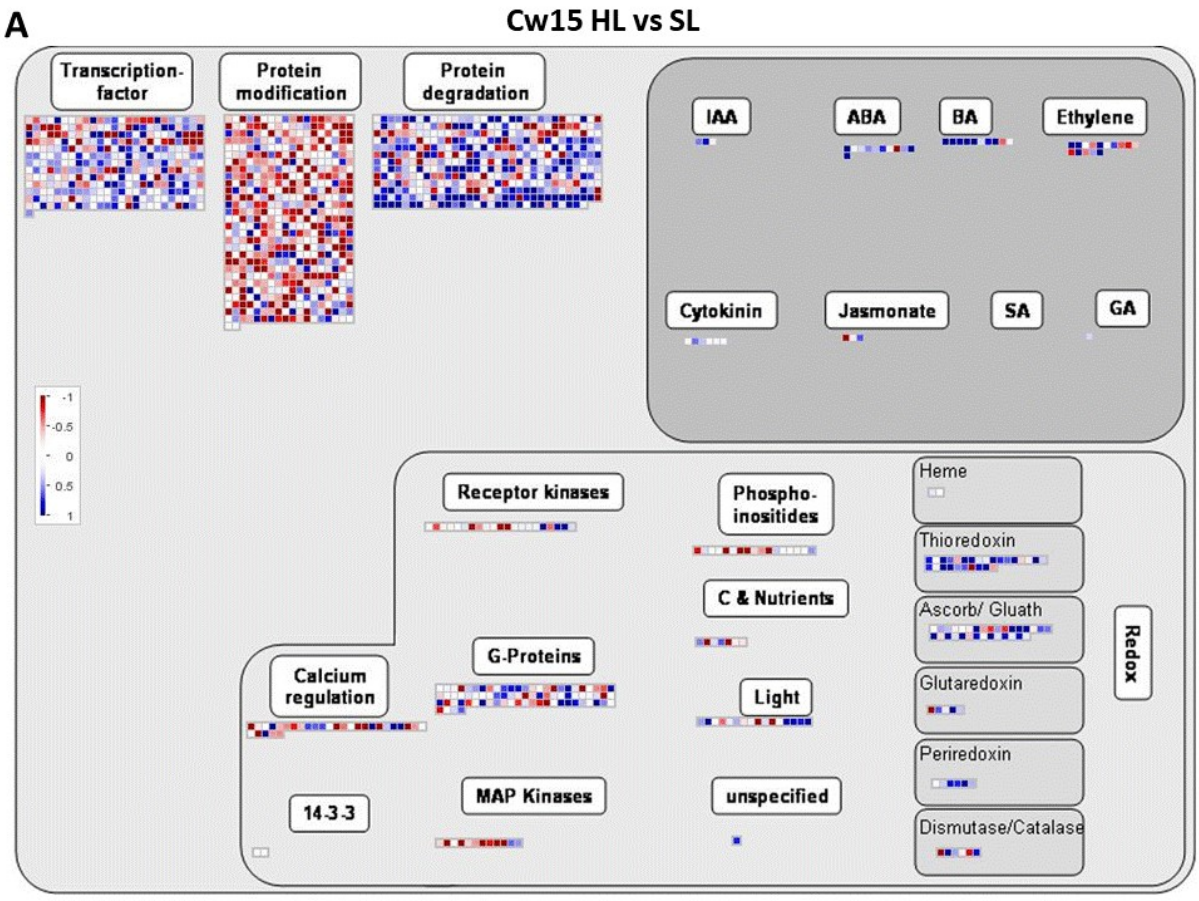
Comparing cw15 to Uex55 at SL levels (figure 6.7D) showed very little difference in transcript levels. In the sucrose degradation pathway two genes encoding sucrose invertases had higher transcript levels in cw15, suggesting that even in standard light levels sucrose breakdown is diminished in the absence of cytochrome c_{6A}. In higher plants invertase can be regulated at the transcriptional level by stress, including abscisic acid signalling (Koch, 2004), therefore there may be an interaction between cytochrome c_{6A} and invertase regulation under high light stress. In the photosynthesis pathway the strongly downregulated gene represents cytochrome c_{6A}, and the two genes with slightly increased transcript levels in Uex55 encode LHCSR1 and ferredoxin, suggesting a possible negative regulation effect of cytochrome c_{6A} on some photosynthesis genes under SL conditions.

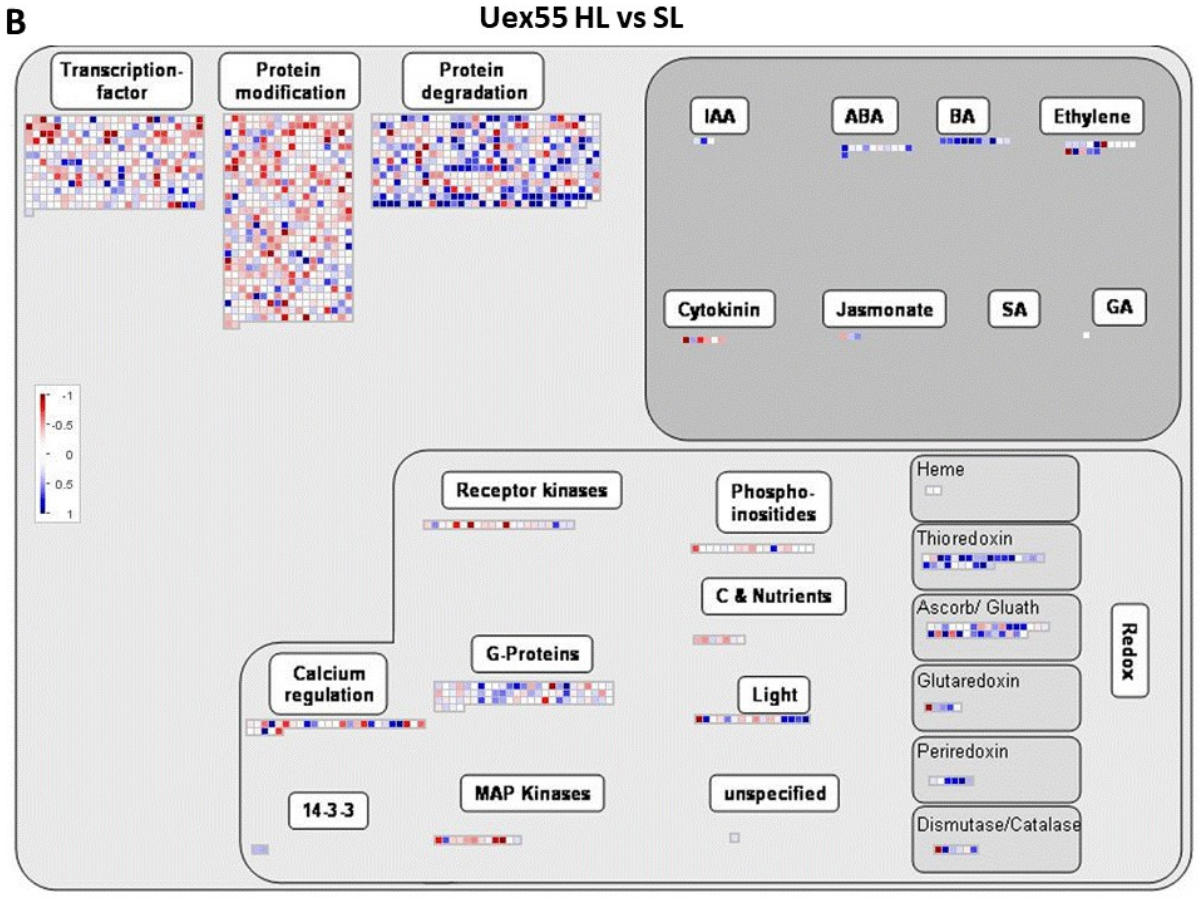
Under HL conditions there were a number of differences in transcript levels between cw15 and Uex55 (figure 6.7C). Tetrapyrrole and isoprenoid synthesis were both found to have higher transcript levels in cw15 than Uex55. This could show regulation of chlorophyll and antioxidant biosynthesis by cytochrome c_{6A}. Proteins involved in photorespiration were also found to have higher transcript levels in cw15 than Uex55, which could also show cytochrome c_{6A} regulating both carbon concentration mechanisms and photorespiration under high light stress, as both mechanisms work to improve the efficiency of carbon fixation. Proteins involved in some catabolic pathways such as the tricarboxylic acid cycle, beta oxidation and amino acid degradation were found to have higher transcript levels in Uex55 than cw15, but looking at the changes in transcript levels for each individual strain under HL stress, this appeared to be due to a lack of downregulation in Uex55 rather than a lack of upregulation in cw15. Therefore once again it appeared that removing cytochrome c_{6A} diminishes overall transcriptional control under HL stress exposure.

6.2.3.3.2. Regulation overview

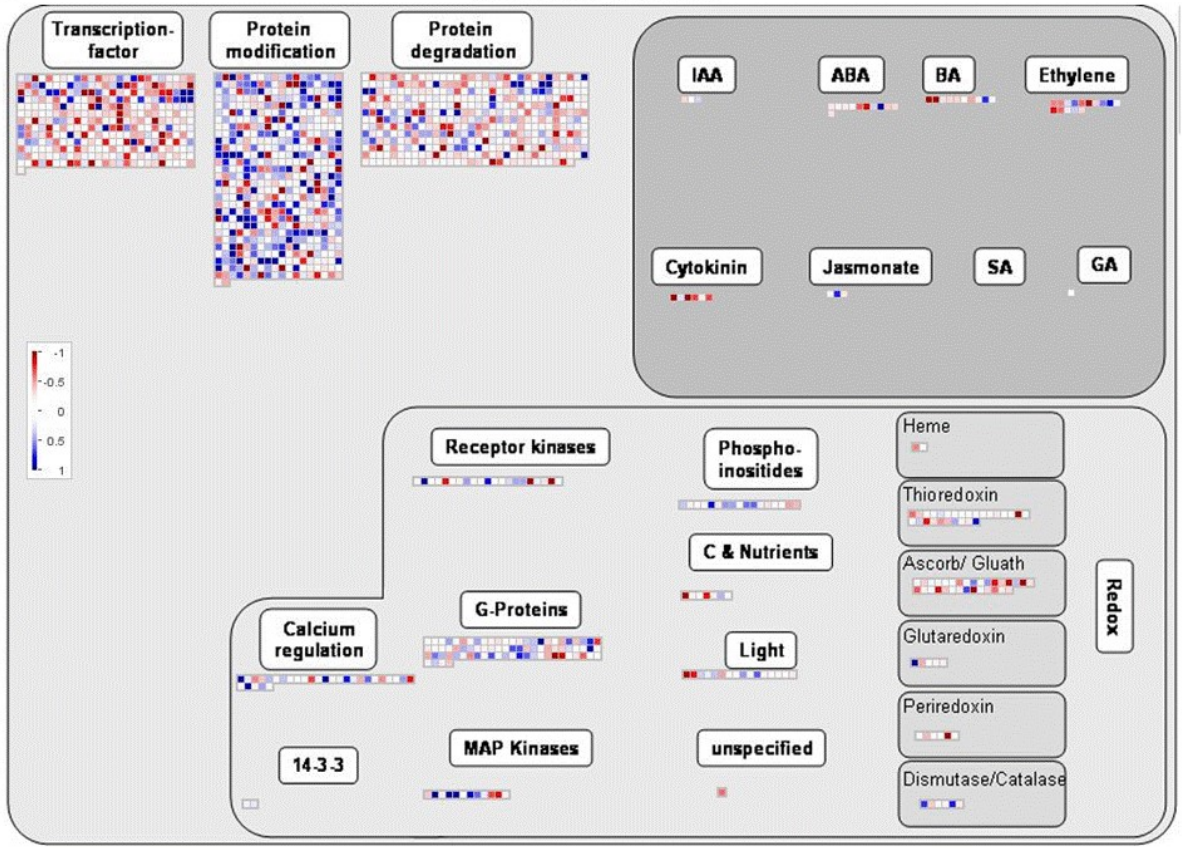
The effect of knocking out cytochrome c_{6A} on regulatory genes under HL stress in *C. reinhardtii* was analysed (figure 6.8). Again, up or downregulation of regulatory gene transcription appers stronger with cytochrome c_{6A} present. Looking at what specific factors in these groups have been altered and what their downstream targets are should be the subject of extensive further study. Genes

associated with glutathione and ascorbate redox regulation and cytokinin signal transduction also had higher transcript levels in cw15 than Uex55. The glutathione-ascorbate cycle is known to be involved in reducing hydrogen peroxide and thus protecting the organism from ROS (figure 1.4; Mendez-Alvarez *et al.*, 1999; Pfannschmidt 2003), therefore this potentially shows a link between cytochrome c_{6A} and the H_2O_2 stress response. Little is known about cytokinin's specific role in *C. reinhardtii*, but there may be overlap between cytokinin and cytochrome c_{6A} signalling (Riaño-Pachón *et al.*, 2008; Žižková *et al.*, 2016). Transcripts associated with other hormone signalling pathways, such as those involving brassinosteroids and abscisic acid, showed upregulation under HL for both strains, although these upregulations were stronger in cw15. Under SL conditions (figure 6.8D) there were no noticeable patterns of difference between Uex55 and cw15.


A
Cw15 HL vs SL
Transcription-factor
Protein modification
Protein degradation
IAA
ABA
BA
Ethylene
Cytokinin
Jasmonate
SA
GA
Receptor kinases
Phospho-inositides
C & Nutrients
G-Proteins
Calcium regulation
Light
14-3-3
MAP Kinases
unspecified
Heme
Thioredoxin
Ascorb/ Gluath
Glutaredoxin
Periredoxin
Dismutase/Catalase
Redox


B
Uex55 HL vs SL
Transcription-factor
Protein modification
Protein degradation
IAA
ABA
BA
Ethylene
Cytokinin
Jasmonate
SA
GA
Receptor kinases
Phospho-inositides
C & Nutrients
G-Proteins
Calcium regulation
Light
14-3-3
MAP Kinases
unspecified
Heme
Thioredoxin
Ascorb/ Gluath
Glutaredoxin
Periredoxin
Dismutase/Catalase
Redox

C


HL Uex55 vs cw15
Transcription-factor
Protein modification
Protein degradation
IAA
ABA
BA
Ethylene
Cytokinin
Jasmonate
SA
GA
-1
-0.5
0
0.5
1
Receptor kinases
Phospho-inositides
C & Nutrients
G-Proteins
Calcium regulation
Light
14-3-3
MAP Kinases
unspecified
Heme
Thioredoxin
Ascorb/ Gluath
Glutaredoxin
Periredoxin
Dismutase/Catalase
Redox

D

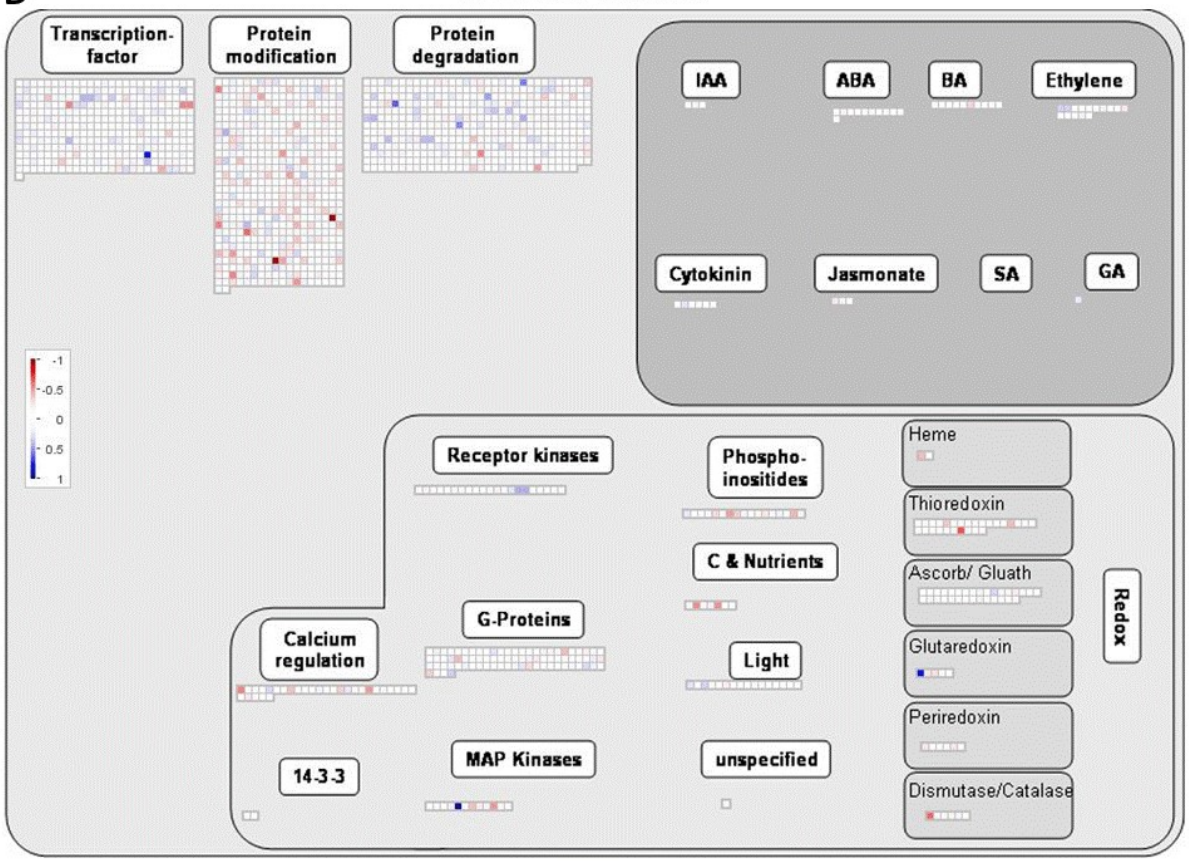
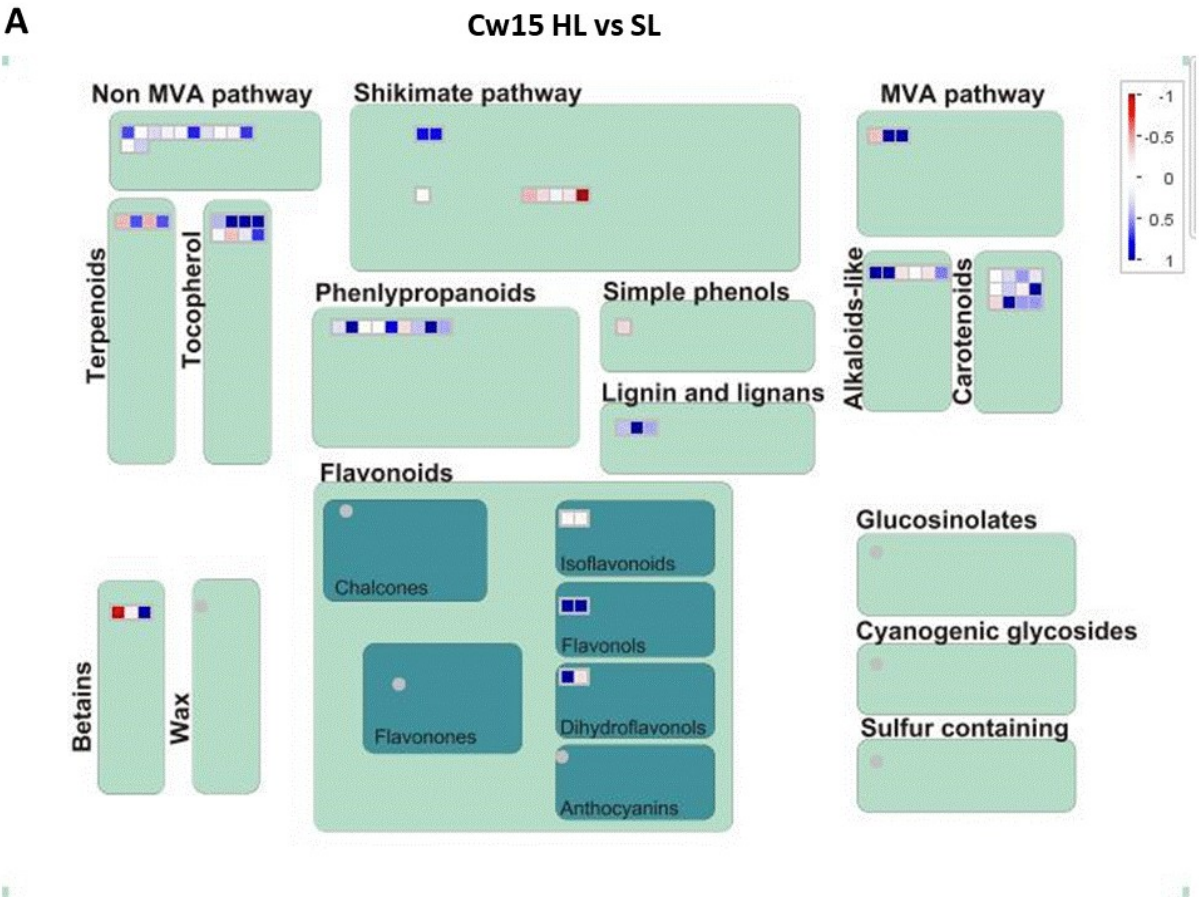

SL Uex55 vs cw15
Transcription-factor
Protein modification
Protein degradation
IAA
ABA
BA
Ethylene
Cytokinin
Jasmonate
SA
GA
-1
-0.5
0
0.5
1
Receptor kinases
Phospho-inositides
C & Nutrients
G-Proteins
Calcium regulation
Light
14-3-3
MAP Kinases
unspecified
Heme
Thioredoxin
Ascorb/ Gluath
Glutaredoxin
Periredoxin
Dismutase/Catalase
Redox

Figure 6.8: A comparison of transcript levels of cw15 and Uex55 under standard (40 µmol.s^{-1}.m^{-2}) and high (400 µmol.s^{-1}.m^{-2}) light conditions in the C. reinhardtii Mapman pathway "Regulation overview". The blue squares represent a higher transcript level under high light stress relative to standard light (A and B) or higher transcript levels in Uex55 compared to cw15 (C and D), whereas red squares represent the opposite.

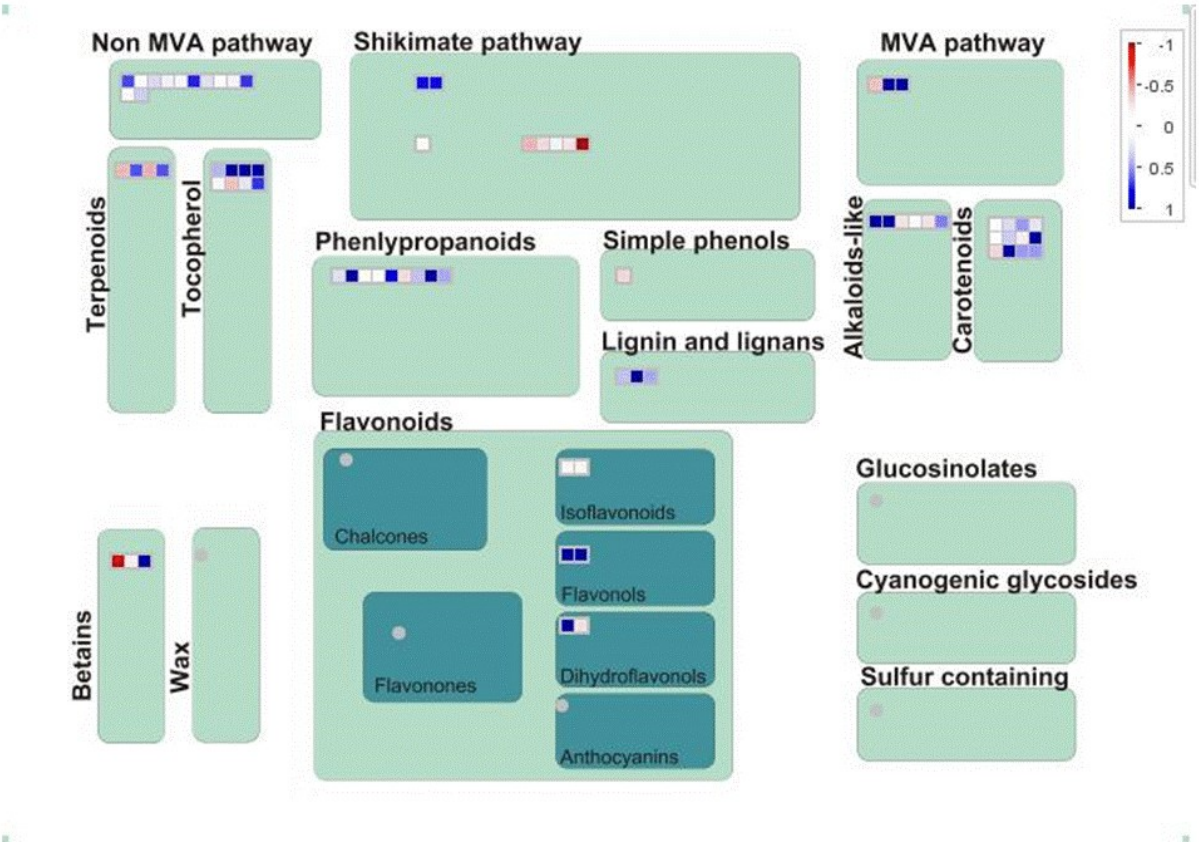
6.2.3.3.3. Secondary metabolism

The effect of HL stress on transcripts associated with secondary metabolism was investigated in cw15 and Uex55 (figure 6.9). Under HL stress, upregulation of transcripts for pathways related to chlorophyll and carotenoid biosynthesis was observed in both cw15 and Uex55 (figure 6.9A and B). This is consistent with the idea that chlorophyll content and antioxidants are increased in the presence of higher irradiance to assist in quenching mechanisms. However, there was a difference in the magnitude of upregulation, with Uex55 demonstrating a diminished upregulation compared to cw15 (figure 6.9C). This was particularly notable in the mevalonate and non-mevalonate pathways, and carotenoid, tocopherol, dihydroflavonol and flavonoid synthesis. Overall this suggests that cytochrome c_{6A} may contribute to the upregulation of pigment and antioxidant synthesis, although the partial elimination of this effect in the knockout line suggested another independent pathway is involved. Again no noticeable difference in transcript levels between Uex55 and cw15 was seen at SL (figure 6.9D).

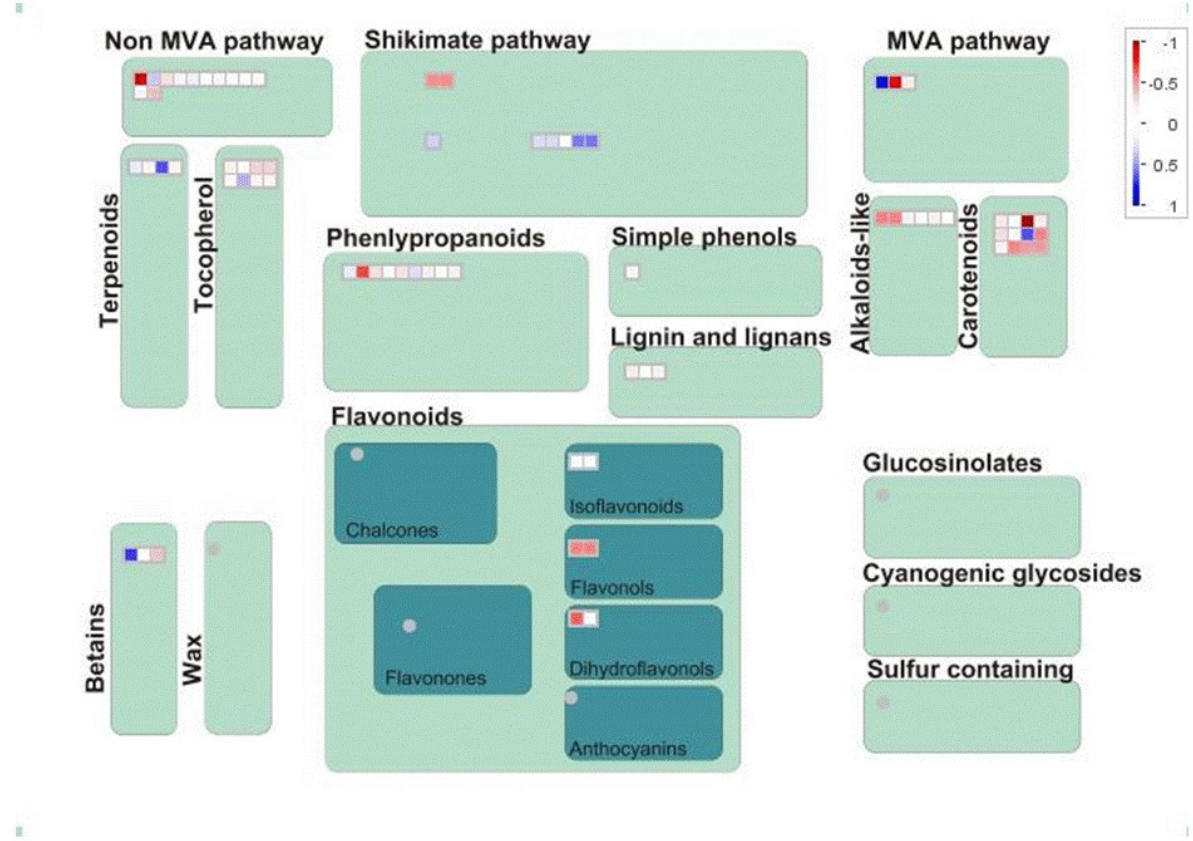
A


Cw15 HL vs SL
Non MVA pathway
Shikimate pathway
MVA pathway
Terpenoids
Tocopherol
Phenlypropanoids
Simple phenols
Alkaloids-like
Carotenoids
Lignin and lignans
Flavonoids
Chalcones
Isoflavonoids
Flavonols
Flavonones
Dihydroflavonols
Anthocyanins
Glucosinolates
Cyanogenic glycosides
Sulfur containing
Betains
Wax
-1
-0.5
0
0.5
1

B


Cw15 HL vs SL
Non MVA pathway
Shikimate pathway
MVA pathway
Terpenoids
Tocopherol
Phenlypropanoids
Simple phenols
Alkaloids-like
Carotenoids
Lignin and lignans
Flavonoids
Chalcones
Isoflavonoids
Flavonols
Flavonones
Dihydroflavonols
Anthocyanins
Glucosinolates
Cyanogenic glycosides
Sulfur containing
Betains
Wax
-1
-0.5
0
0.5
1

C


HL Uex55 vs cw15
Non MVA pathway
Shikimate pathway
MVA pathway
Terpenoids
Tocopherol
Phenlypropanoids
Simple phenols
Alkaloids-like
Carotenoids
Lignin and lignans
Flavonoids
Chalcones
Isoflavonoids
Flavonols
Flavonones
Dihydroflavonols
Anthocyanins
Glucosinolates
Cyanogenic glycosides
Sulfur containing
Betains
Wax
-1
-0.5
0
0.5
1

D

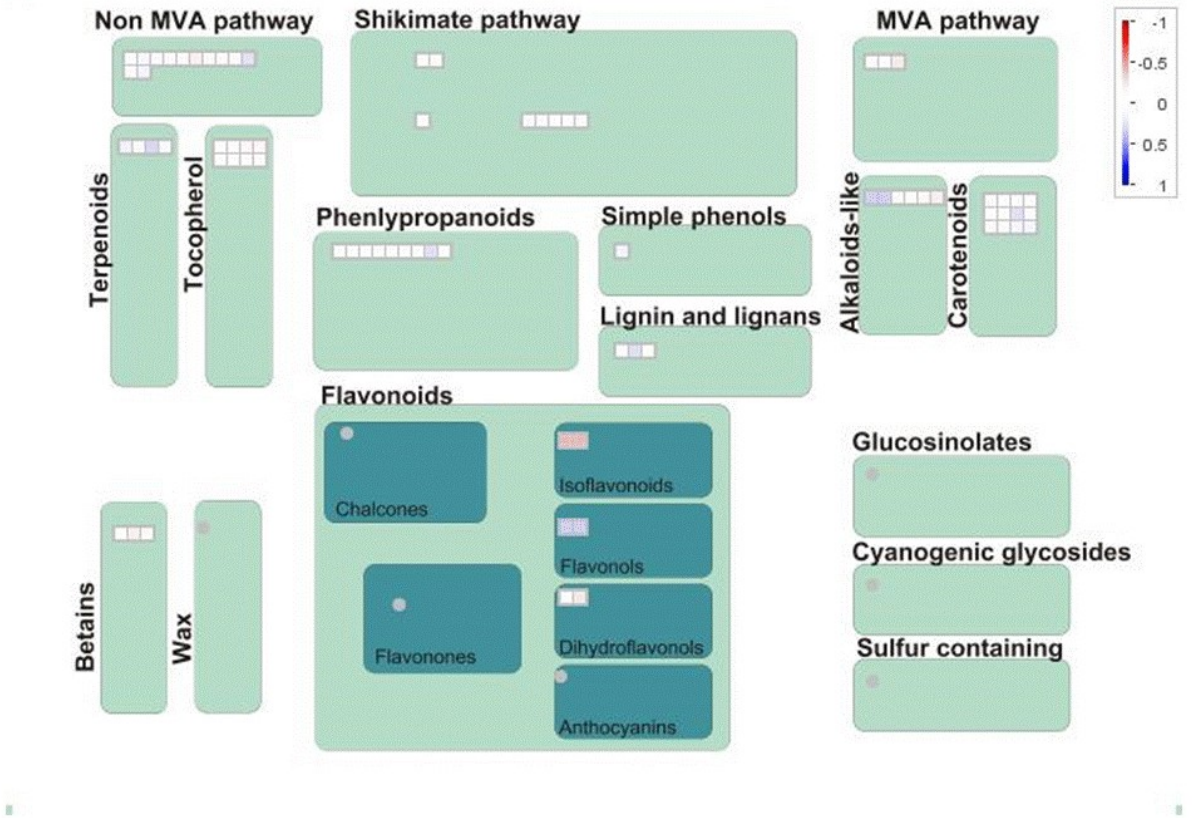
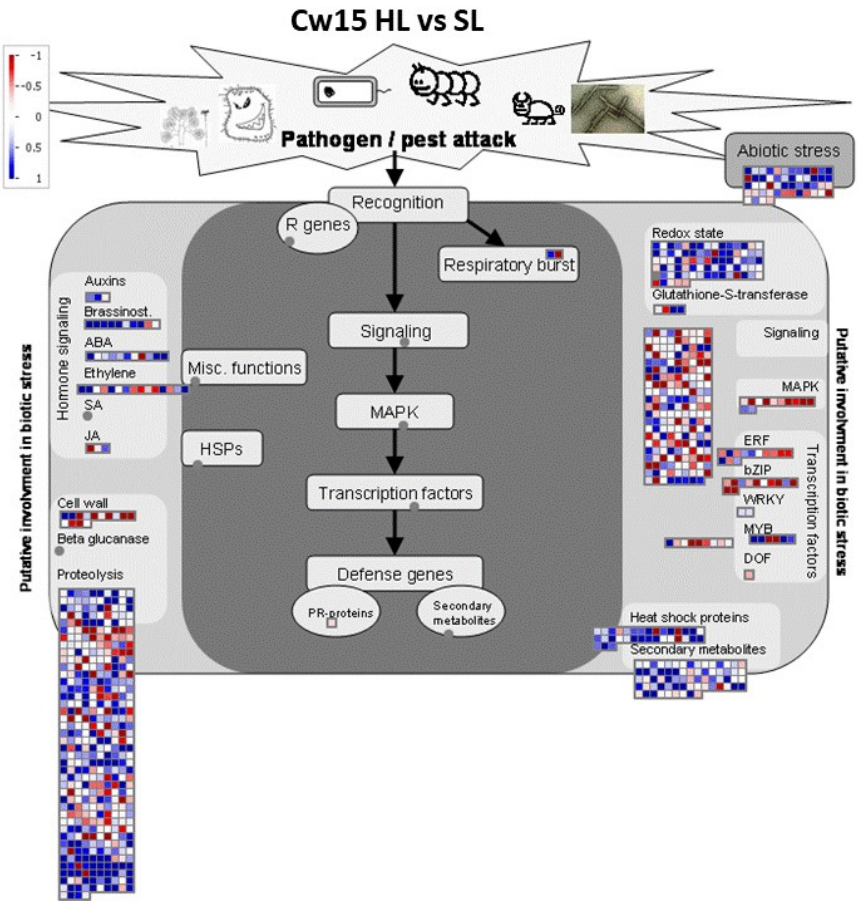

SL Uex55 vs cw15
Non MVA pathway
Shikimate pathway
MVA pathway
Terpenoids
Tocopherol
Phenlypropanoids
Simple phenols
Alkaloids-like
Carotenoids
Lignin and lignans
Flavonoids
Chalcones
Isoflavonoids
Flavonols
Flavonones
Dihydroflavonols
Anthocyanins
Glucosinolates
Cyanogenic glycosides
Sulfur containing
Betains
Wax
-1
-0.5
0
0.5
1

Figure 6.9: A comparison of transcript levels of cw15 and Uex55 under standard (40 µmol.s^{-1}.m^{-2}) and high (400 µmol.s^{-1}.m^{-2}) light conditions in the C. reinhardtii Mapman pathway "Secondary Metabolism". The blue squares represent a higher transcript level under high light stress relative to standard light (A and B) or higher transcript levels in Uex55 compared to cw15 (C and D), whereas red squares represent the opposite.

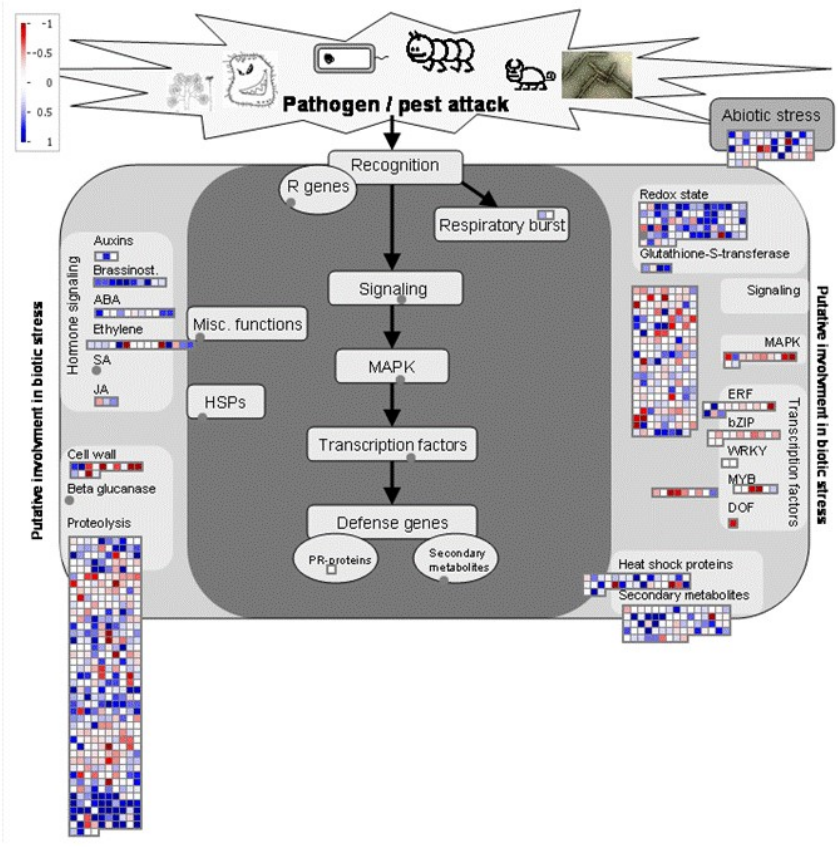
6.2.3.3.4. Biotic and abiotic stress

The relative transcript levels of biotic and abiotic stress genes were analysed under HL stress in cw15 and Uex55 (figure 6.10). As noted in figure 6.7, thioredoxin metabolism and the ascorbate glutathione cycle had higher transcript levels in cw15 than Uex55 under HL conditions. Heat stress and heat shock genes also had notably higher transcript levels in cw15 upon HL stress than in Uex55. As an increase in irradiance usually leads to a higher temperature, it would be expected that heat shock genes would be upregulated upon HL stress. The diminished heat shock transcript upregulation in Uex55 suggests that cytochrome c_{6A} is involved in the activation of the heat shock response. Many signalling gene transcripts appeared to have differential regulation in cw15 and Uex55, particularly MAP kinases, transcription factors and genes for proteins involved in proteolysis. Although again the precise targets of these genes require investigation, it is useful to note that knocking out cytochrome c_{6A} appears to have a large effect on *C. reinhardtii* regulatory pathways, which could suggest either a direct signalling role for cytochrome c_{6A} or that without cytochrome c_{6A} the cell needs to adjust to HL stress differently. Under SL conditions two genes showed a large difference in transcript level between cw15 and Uex55. One was a MAP kinase, which had higher transcript levels in Uex55 than cw15, and the other was a potential 2-oxoglutarate-dependent dioxygenase (2-ODD). The 2-ODD family of enzymes is very broad, and is involved in oxygenation reactions, usually within biosynthetic processes (Kawai *et al.*, 2014). Therefore some genes involved in metabolism or regulation do appear to be regulated directly or indirectly by cytochrome c_{6A} under SL conditions, but no clear pattern can be determined from them.

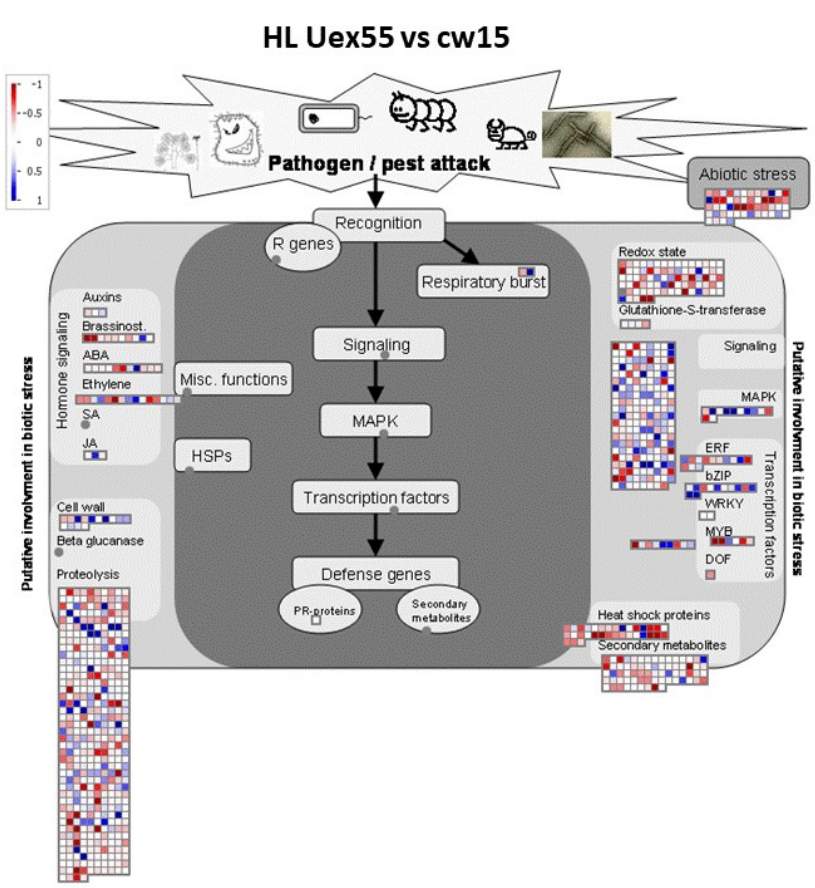
A


Cw15 HL vs SL
-1
-0.5
0
0.5
1
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Redox state
Glutathione-S-transferase
Signaling
Misc. functions
MAPK
HSPs
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Heat shock proteins
Secondary metabolites
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Putative involvment in biotic stress
Putative involvment in biotic stress
Signaling
MAPK
ERF
bZIP
WRKY
MYB
DOF
Transcription factors

B


Uex55 HL vs SL
-1
-0.5
0
0.5
1
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Redox state
Glutathione-S-transferase
Signaling
Misc. functions
MAPK
HSPs
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Heat shock proteins
Secondary metabolites
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Putative involvment in biotic stress
Putative involvment in biotic stress
Signaling
MAPK
ERF
bZIP
WRKY
MYB
DOF
Transcription factors

C


HL Uex55 vs cw15
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Signaling
MAPK
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Misc. functions
HSPs
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Redox state
Glutathione-S-transferase
Signaling
MAPK
Transcription factors
ERF
bZIP
WRKY
MYB
DOF
Heat shock proteins
Secondary metabolites
Putative involvment in biotic stress
Putative involvment in biotic stress

D

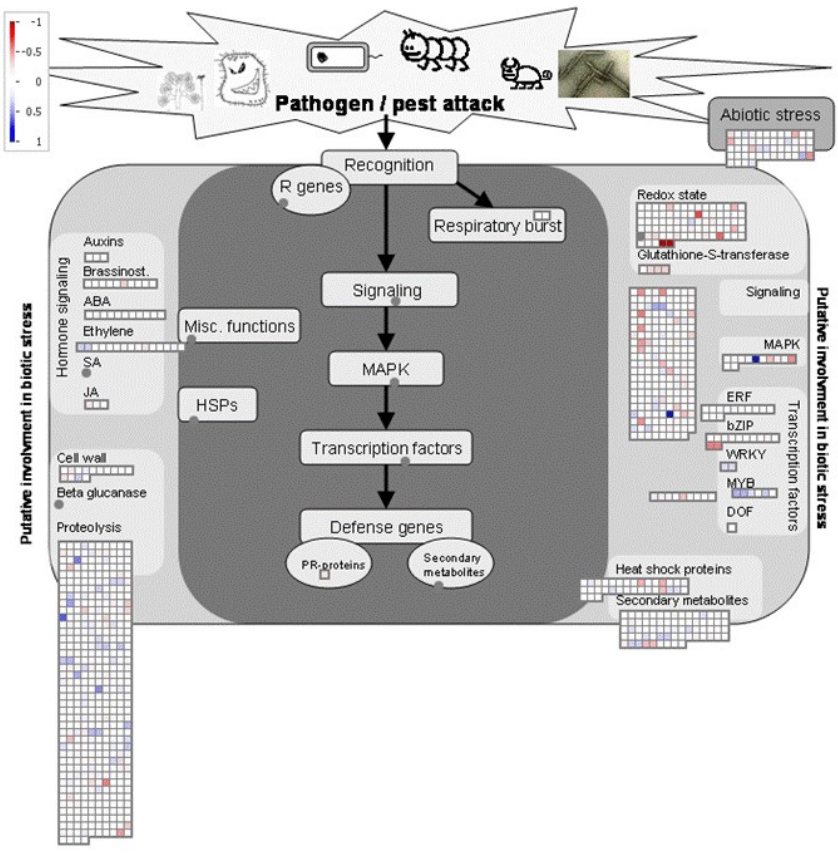
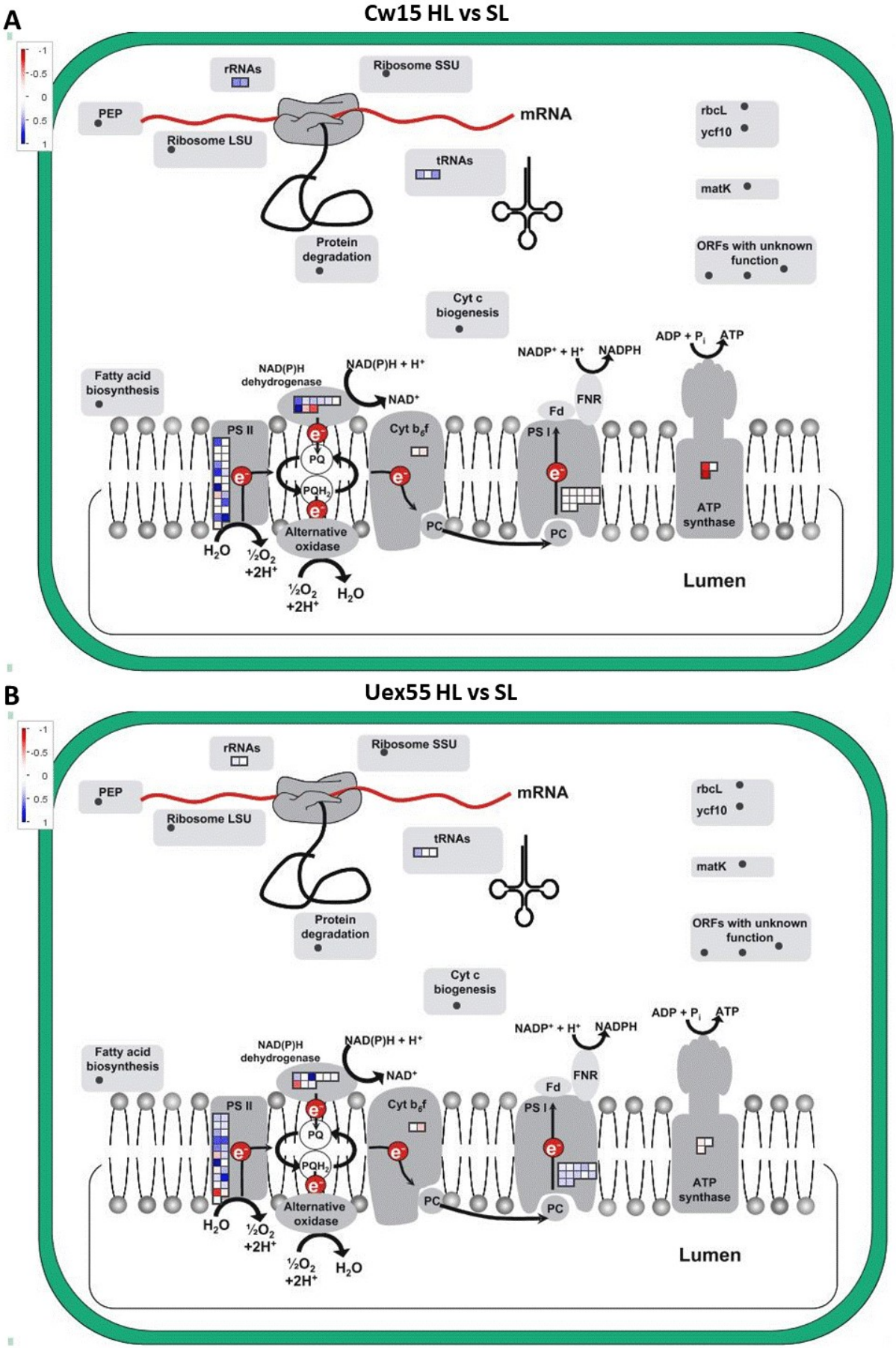

SL Uex55 vs cw15
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Signaling
MAPK
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Misc. functions
HSPs
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Redox state
Glutathione-S-transferase
Signaling
MAPK
Transcription factors
ERF
bZIP
WRKY
MYB
DOF
Heat shock proteins
Secondary metabolites
Putative involvment in biotic stress
Putative involvment in biotic stress

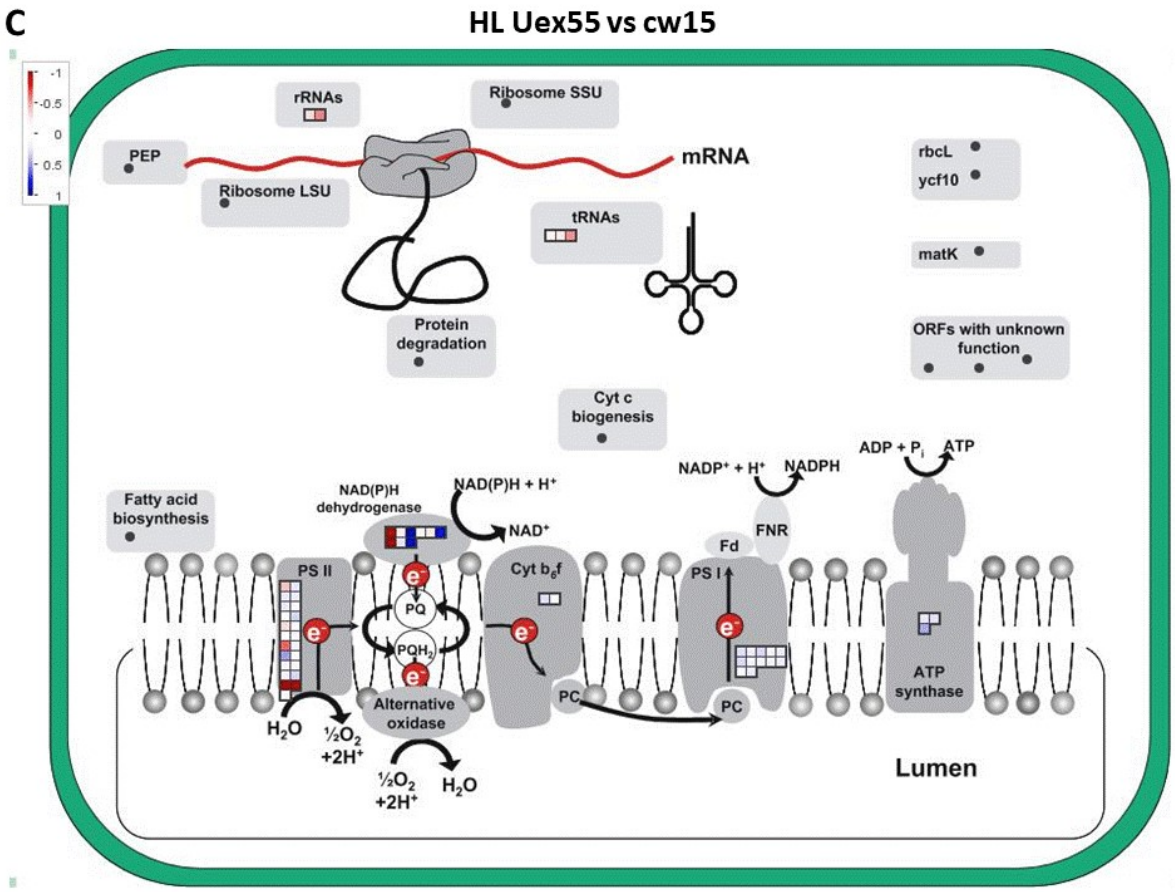
Figure 6.10: A comparison of transcript levels of cw15 and Uex55 under standard (40 µmol.s^{-1}.m^{-2}) and high (400 µmol.s^{-1}.m^{-2}) light conditions in the C. reinhardtii Mapman pathway "Biotic stress". The blue squares represent a higher transcript level under high light stress relative to standard light (A and B) or higher transcript levels in Uex55 compared to cw15 (C and D), whereas red squares represent the opposite.

6.2.3.3.5. Chloroplast overview

The changes in transcript level under HL stress for cw15 and Uex55 were finally analysed for chloroplast-targeted genes, particularly focussing on photosynthetic electron transport components (figure 6.11). At SL the only difference seen between Uex55 and cw15 (figure 6.11D) was the box representing cytochrome c_{6A} transcripts. Under HL stress in both cw15 and Uex55 (figure 6.11A and B) there was an increase in transcript level of genes encoding photosystem II components. However, maturation factors for *psbB* and *psbD* mRNA both showed an upregulation in cw15 under HL that was not seen in Uex55 (figure 6.11C). This might suggest both PSII assembly and repair are partially under regulation by cytochrome c_{6A}, or that the consequences of cytochrome c_{6A} loss indirectly result in the need for more psbB and psbD synthesis.


A
Cw15 HL vs SL
-1
-0.5
0
0.5
1
rRNAs
Ribosome SSU
PEP
mRNA
Ribosome LSU
tRNAs
rbcL
ycf10
matK
Protein degradation
ORFs with unknown function
Cyt c biogenesis
ADP + Pi ATP
NADP+ + H+ NADPH
Fatty acid biosynthesis
NAD(P)H dehydrogenase
NAD(P)H + H+
NAD+
FNR
Fd
PS II
Cyt b6f
PS I
PQ
PQH2
ATP synthase
Alternative oxidase
PC
H2O
½O2 +2H+
Lumen
B
Uex55 HL vs SL
-1
-0.5
0
0.5
1
rRNAs
Ribosome SSU
PEP
mRNA
Ribosome LSU
tRNAs
rbcL
ycf10
matK
Protein degradation
ORFs with unknown function
Cyt c biogenesis
ADP + Pi ATP
NADP+ + H+ NADPH
Fatty acid biosynthesis
NAD(P)H dehydrogenase
NAD(P)H + H+
NAD+
FNR
Fd
PS II
Cyt b6f
PS I
PQ
PQH2
ATP synthase
Alternative oxidase
PC
H2O
½O2 +2H+
Lumen

C


HL Uex55 vs cw15
-1
-0.5
0
0.5
1
rRNAs
Ribosome SSU
PEP
mRNA
Ribosome LSU
tRNAs
rbcL
ycf10
matK
Protein degradation
ORFs with unknown function
Cyt c biogenesis
ADP + Pi
ATP
NADP+ + H+
NADPH
Fatty acid biosynthesis
NAD(P)H dehydrogenase
NAD(P)H + H+
NAD+
FNR
Fd
PS II
Cyt b6f
PS I
PQ
PQH2
ATP synthase
PC
Alternative oxidase
H2O
½O2 +2H+
Lumen

D

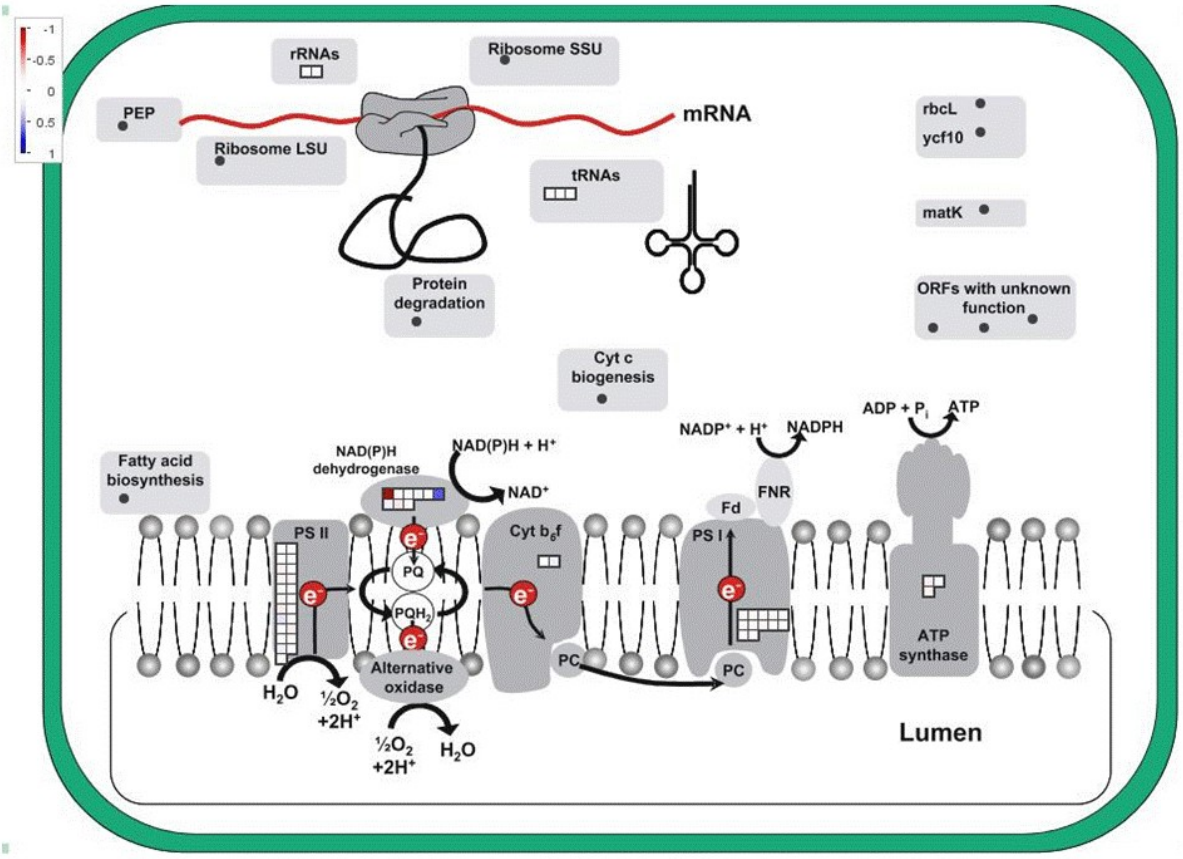

SL Uex55 vs cw15
-1
-0.5
0
0.5
1
rRNAs
Ribosome SSU
PEP
mRNA
Ribosome LSU
tRNAs
rbcL
ycf10
matK
Protein degradation
ORFs with unknown function
Cyt c biogenesis
ADP + Pi
ATP
NADP+ + H+
NADPH
Fatty acid biosynthesis
NAD(P)H dehydrogenase
NAD(P)H + H+
NAD+
FNR
Fd
PS II
Cyt b6f
PS I
PQ
PQH2
ATP synthase
PC
Alternative oxidase
H2O
½O2 +2H+
Lumen

Figure 6.11: A comparison of transcript levels of cw15 and Uex55 under standard (40 µmol.s^{-1}.m^{-2}) and high (400 µmol.s^{-1}.m^{-2}) light conditions in the C. reinhardtii Mapman custom pathway "Chloroplast". The blue squares represent a higher transcript level under high light stress relative to standard light (A and B) or higher transcript levels in Uex55 compared to cw15 (C and D), whereas red squares represent the opposite.

The other notable differences in transcript level between cw15 and Uex55 under HL stress were for the ferredoxin isoforms. FDX1/PetF was found to be upregulated in cw15 under HL stress to a greater extent than in Uex55, but FDX2 and FSX5 had higher transcript levels under HL stress in Uex55 than cw15. FDX6 and FDX4 showed no notable difference between the two strains. FDX1 is known to be the main isoform of ferredoxin, in the sense that it is represented by 98 % of the total ferredoxin transcripts (Terauchi *et al*., 2009; Winkler *et al*., 2010). Therefore the upregulation of FDX1 transcripts under HL appears logical as it might be associated with both linear and cyclic electron transport. The fact that in Uex55 this upregulation was diminished again shows a potential regulatory role for cytochrome c_{6A} in linear and cyclic photosynthetic electron transport under HL stress. FDX2 has been shown to be induced when nitrate is present instead of ammonia (Terauchi *et al.*, 2009), which could link to the upregulation of nitrate transporters noted in the top ten most upregulated genes in Uex55 under HL exposure relative to cw15 (table 6.3), suggesting that without cytochrome c_{6A} HL induces nitrate transport into the cell. FDX5 has been shown to be induced under anaerobic conditions and copper replete conditions (Jacobs *et al.*, 2009), thus suggesting that a lack of cytochrome c_{6A} causes or simulates either of these effects on the cells.

Overall the Mapman analysis supported the observation that cytochrome c_{6A} appears essential for the regulation of certain gene transcripts under HL stress, such as pigment synthesis, gene regulation and biotic stress responses, and lends support for a signalling hypothesis.

6.2.3.4. Looking at changes in transcript level of key stress markers in C. reinhardtii

Past studies on the transcriptomic responses to singlet oxygen, hydrogen peroxide and HL stress in *C. reinhardtii* have identified key genes known to be regulated (Urzica *et al.*, 2012; Mettler *et al.*, 2014; Wakao *et al.*, 2014). Using these stress markers the interaction between cytochrome c_{6A} and these stress response pathways was studied identified.

6.2.3.4.1. Singlet oxygen stress

Genes whose transcript level increased upon singlet oxygen exposure in the study by Wakao *et al*. (2014) were investigated in the RNA sequencing data (figure 6.12). For a majority of the genes showing a strong difference in regulation under HL stress between the two strains, cw15 showed a strong upregulation of transcripts that was diminished in Uex55. The most obvious exception to this was *GST1*, a glutathione-S-transferase which showed a mildly stronger upregulation in Uex55 under HL stress relative to cw15. If this change were significant, it could show a specific oxidative stress induced by an absence of cytochrome c_{6A}.

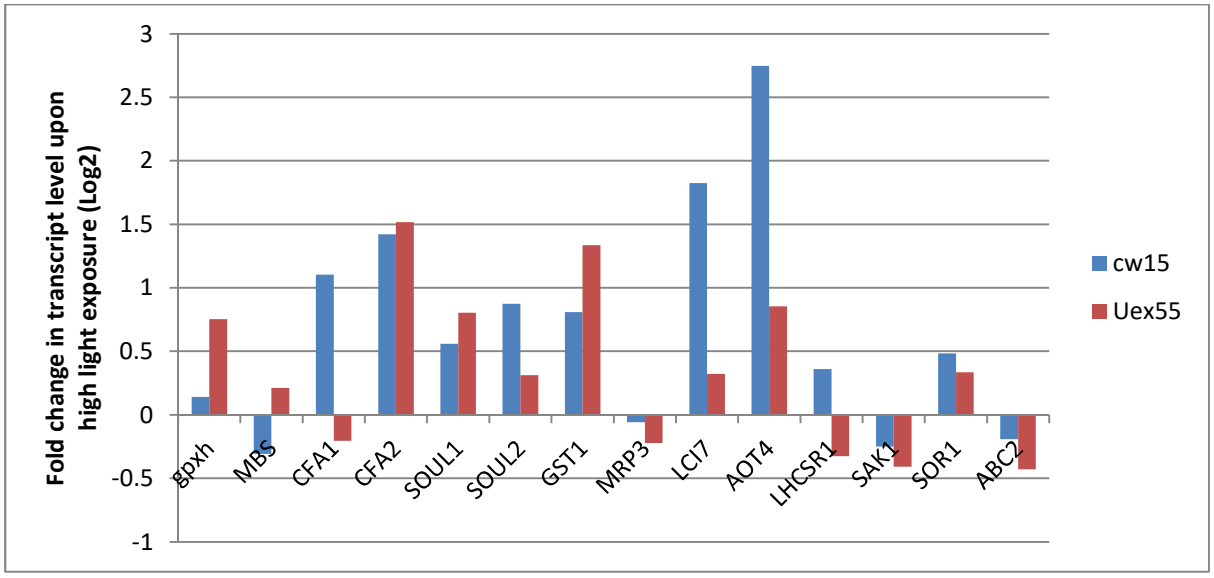


Figure 6.12: Comparison of transcript level change upon high light exposure for 4 h in cw15 (blue) and Uex55 (red) for genes identified as upregulated under singlet oxygen exposure by Wakao et al. (2014).

Three genes demonstrated to be upregulated by singlet oxygen stress also appeared to be regulated at least in part by cytochrome c_{6A}. *LCI7* is a gene that links together carbon concentration, hydrogen peroxidase reduction and the singlet oxygen response. *LCI7* has been shown to be upregulated under low CO_2 conditions (Yamano *et al.*, 2008), and on Phytozome is annotated as encoding an alkyl hydroperoxide reductase C-related, which is an enzyme that reduces hydrogen peroxide in bacteria (Seaver and Imlay, 2001). *AOT4* encodes a putative cysteine transporter that has been found to be upregulated in response to sulphur deprivation (Blaby *et al.*, 2015), therefore could be contributing to cysteine synthesis under singlet oxygen or HL stress by acting as a sink for reduction to allow for increased quenching. *CFA1* encodes a cyclopropane fatty acid synthase, which has been implied in the oxidative stress response in bacteria (Kim *et al.*, 2005). The function of *CFA1* in *C. reinhardtii* is not clear, but it has been found in the oil body proteome (Nguyen *et al.*, 2011), and therefore could

be involved in increased synthesis again to act as a sink for ATP and reducing equivalents under HL stress. Therefore there appears to be some overlap in cytochrome c_{6A} and singlet oxygen transcriptional regulation, but these observations may be from an indirect effect.

6.2.3.4.2. Hydrogen peroxide stress

Transcripts for genes shown to be regulated in response to hydrogen peroxide stress by Urzica *et al.* (2012) were analysed in the RNA sequencing data (figure 6.13). A majority of the selected genes did not show a noticeable difference in transcript levels between cw15 and Uex55 upon HL stress except *MSD3*. *MSD3*, which encodes a superoxide dismutase, was downregulated in both strains upon HL exposure, but the downregulation was greater in cw15 than Uex55. This appears to conflict with prior observations in the RNAseq data where glutathione ascorbate cycle genes and a putative alkyl hydroperoxide reductase appeared upregulated under HL stress when cytochrome c_{6A} was present. This would suggest that MSD3 may be involved in the hydrogen peroxide stress response in a different context from the upregulated genes, for example using a different electron donor from those used in the glutathione ascorbate cycle. Overall the data suggest that cytochrome c_{6A} does not appear to directly regulate hydrogen peroxide stress related transcripts under HL stress.

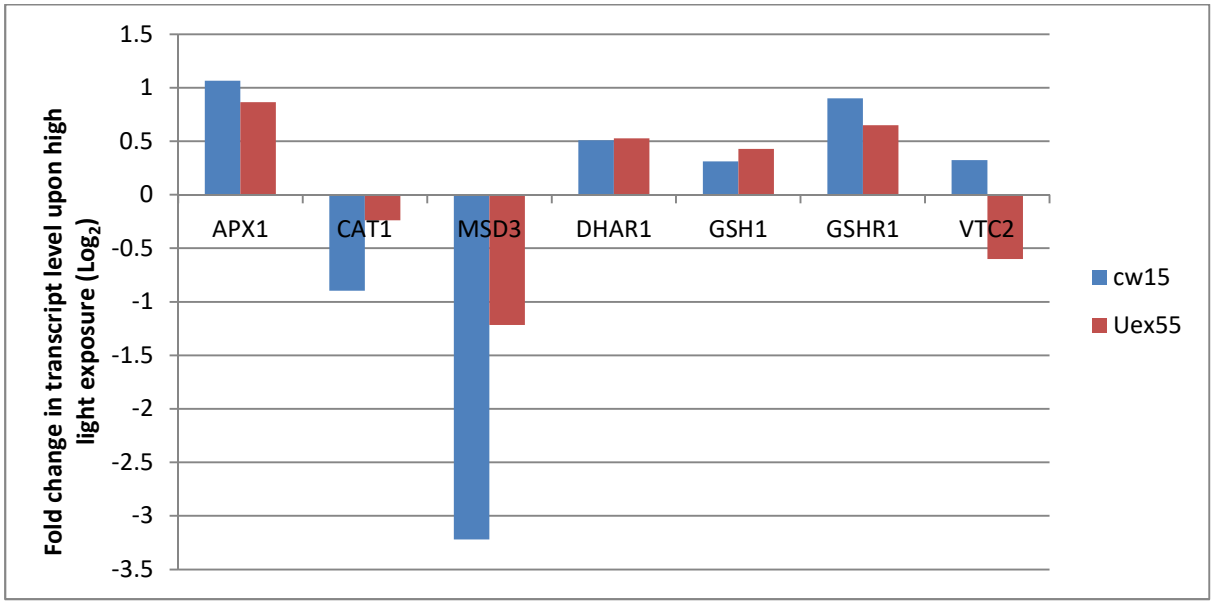


Figure 6.13: Comparison of transcript level change upon high light exposure for 4 h in cw15 (blue) and Uex55 (red) for genes identified as upregulated under hydrogen peroxide exposure in Urzica et al. (2012).

6.2.3.4.3. High light stress

Transcripts of genes shown to be regulated in response to HL stress by Mettler *et al.* (2014) were analysed in the RNA sequencing data (figure 6.14). The CCM genes *CCP2* and *LCI1* were shown here to be significantly upregulated in HL stress in cw15 but not Uex55 cells, indicating a link between CCM, HL and cytochrome c_{6A}. Genes encoding both isoforms of LHCSR3 and CVDE also demonstrated upregulation in cw15 upon HL exposure that was diminished in Uex55. LHCSR3 is a xanthophyll-associated protein involved in short-term NPQ through detecting low luminal pH and responding by quenching excited chlorophyll to prevent singlet oxygen production (Bonente *et al.*, 2011), and CVDE is a chlorophycean-specific form of violaxanthin de-epoxidase, a key enzyme in producing xanthophyll in the xanthophyll cycle (Li *et al.*, 2016). Therefore cytochrome c_{6A} could be involved in regulating NPQ genes in response to HL exposure, which is consistent with the chlorophyll fluorescence data in this study. *HEP1* was also found to have been upregulated upon HL exposure in cw15, but not in Uex55. *HEP1* encodes an import protein that is thought to be involved in escorting the heat shock protein HSP70 into the mitochondria (Kluth *et al.*, 2012). This provides a potential link between heat stress, light stress and cytochrome c_{6A} in *C. reinhardtii*. The *GUN4* gene is a magnesium chelatase that can stimulate chlorophyll biosynthesis and is involved in retrograde signalling from plastid to nucleus (Brzezowski *et al.*, 2014). *ADC1* showed a mild difference in regulation in cw15 under HL stress compared to Uex55, and therefore may be a gene dependent on cytochrome c_{6A} presence. *ADC1* encodes an aminodeoxychorismate synthase, an enzyme important for folate synthesis, which in turn is a cofactor required for synthesis of further metabolites such as amino acids (Gorelova *et al.*, 2019). Therefore this could be another energy sink upregulated under HL stress to help drive photosynthesis.

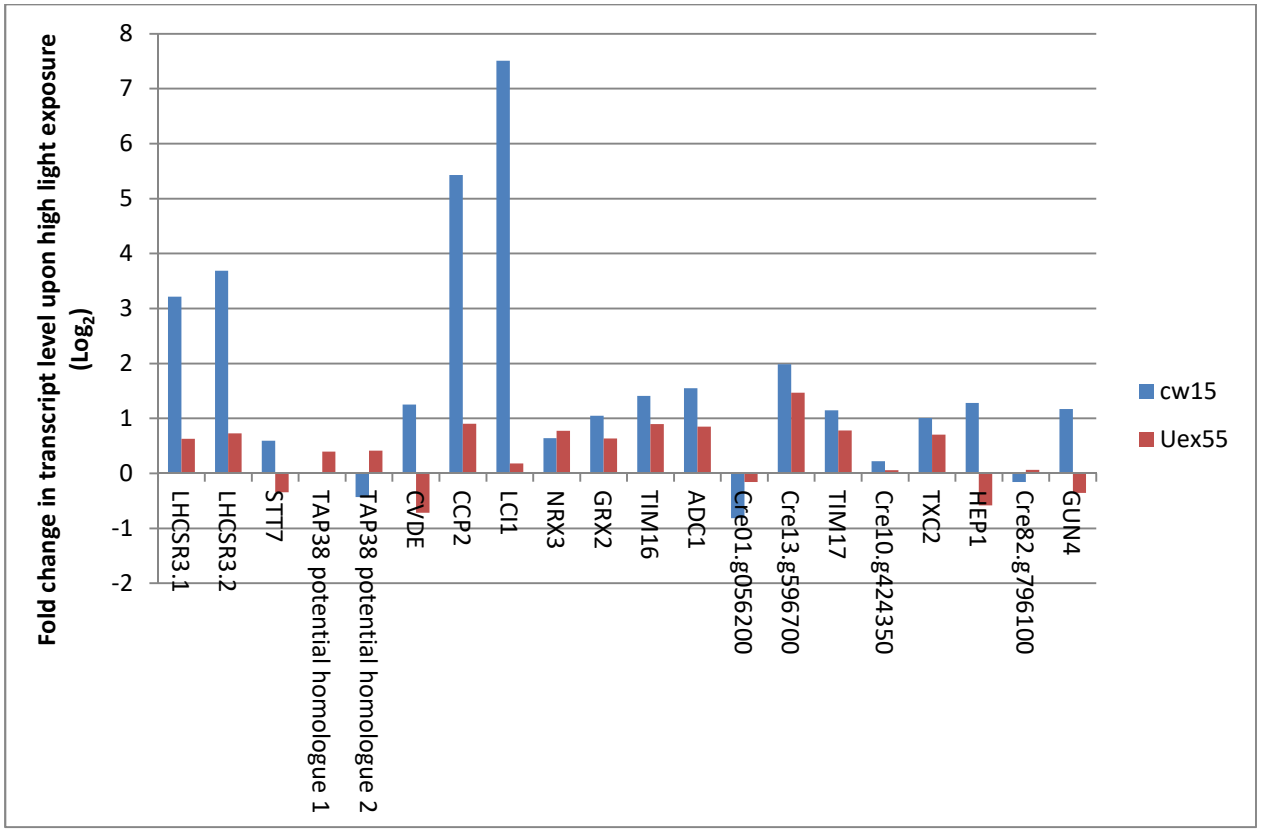


Figure 6.14: Comparison of transcript level change upon high light exposure for 4 h in cw15 (blue) and Uex55 (red) for genes identified as upregulated under high light stress in Mettler et al. (2014). TAP38 potential homologues were located through a pBLAST search using TAP38 from Arabidopsis thaliana.

The kinase STT7 and a phosphorylase currently unidentified in *C. reinhardtii*, but designated TAP38 in *Arabidopsis thaliana*, are responsible for state transition control (Depège *et al.*, 2003; Dumas *et al.*, 2017). Therefore to see if transcripts for proteins involved in state transitions were regulated by cytochrome c_{6A} at the transcript level STT7 and two potential TAP38 homologues (identified through pBLAST of the *A. thaliana* TAP38 peptide sequence) were included in figure 6.14. STT7 demonstrated a low-magnitude upregulation in transcript level under HL stress in cw15, which was reversed in Uex55. Although these values were very low, the data could show a mild cytochrome c_{6A}-dependent regulation of STT7 under HL conditions. The first TAP38 potential homologue showed no changes in transcript levels upon HL stress, but the second potential homologue demonstrated the opposite transcriptional profile from STT7, with a mild upregulation in Uex55 and a mild downregulation in cw15 of transcript levels upon HL exposure. As STT7 and TAP38 functions are opposite and STT7 activity would be favoured under HL stress, the second TAP38 homologue found could be a potential candidate for the state transition phosphorylase in *C. reinhardtii*.

Overall the data suggest a link between cytochrome c_{6A} and the HL stress response as outlined by Mettler *et al.* (2014). However there is not a perfect overlap between cytochrome c_{6A}-HL regulated genes and the genes upregulated in the study from Mettler *et al.* (2014), which could show that there are aspects of the HL stress response independent of cytochrome c_{6A}. It should also be noted, however, that the HL conditions of the two studies being compared were not the same, as the study described here used 400 µE m^{-2} s^{-1} and the study of Mettler et al. used 145 µE m^{-2} s^{-1}.

6.3. Discussion

6.3.1. Implications for the function of cytochrome c_{6A} in the high light stress response

6.3.1.1. Cytochrome c_{6A} may be involved in retrograde signalling in response to stress

A key observation of the RNA sequencing data was that most of the differences in transcriptional regulation upon HL exposure between Uex55 and cw15 were the results of a significant change in cw15 that was diminished in Uex55. This suggests that when cytochrome c_{6A} is absent from *C. reinhardtii*, certain signalling pathways that lead to transcriptional regulation in the nucleus in response to HL failed to be activated. This overall provides more evidence towards a signalling role for cytochrome c_{6A} under HL stress than as an alternative electron route, as the absence of a safety valve would be likely to have resulted in an increase in stress response gene transcript levels upon HL stress. It was also observed that cytochrome c_{6A} transcript levels did not increase greatly upon HL exposure, which, combined with the observation that cytochrome c_{6A} has a low expression level under standard light (Nimmo, 2011), again provides evidence against a safety valve model, under which one might expect that the protein level would need to increase to have a significant effect on electron transport. There are multiple retrograde signalling pathways in a complex overlapping network involving singlet oxygen, the redox poise of the plastoquinone pool, and light detection at the eyespot (Teramoto *et al.*, 2002; Wakao *et al.*, 2014; Erickson *et al.*, 2015). Therefore under HL conditions cytochrome c_{6A} could be detecting an over-reduced plastoquinol pool through becoming reduced itself, and then passing that electron to a downstream signalling molecule to convey a HL stress signal to the nucleus.

Two groups of genes that cytochrome c_{6A} may upregulate are these encoding proteins involved in carotenoid biosynthesis and CCM. The carotenoid biosynthesis upregulation is consistent with both the growth curve results from chapter 4 under fluctuating light and the dark to light transition (in which cytochrome c_{6A} knockout mutants displayed diminished growth under these light conditions), and the results from chlorophyll fluorescence (in which cytochrome c_{6A} knockout mutants displayed

a decreased NPQ response). These observations support a link between cytochrome c_{6A} and NPQ. This is also supported by the effects of *LHCSR3.1* and *LHCSR3.2* transcript levels under HL exposure by cytochrome c_{6A}. However, the chlorophyll fluorescence studies were performed over too short a timescale for transcriptional change to be involved, and the cultures were grown beforehand at SL intensity prior to the experiment, so there may also be a more direct role of cytochrome c_{6A} in NPQ in addition to longer term effects. More chlorophyll fluorescence studies on HL adapted cultures could provide more information regarding the relationship between cytochrome c_{6A} and NPQ. Cytochrome c_{6A} involvement in CCM upregulation under HL stress could provide information on the link between photosynthesis and CCM, and is supported by the cytochrome c_{6A}-controlled upregulation of photorespiration. The overlap between cytochrome c_{6A}-affected transcripts and CIA5-affected CCM transcripts was strong, perhaps suggesting an interaction of cytochrome c_{6A} with the transcriptional regulator in retrograde signalling. However it must be noted that CIA5 regulates a much larger range of gene transcripts than just CCM, and so cytochrome c_{6A} may interact with the CIA5-regulatory pathway, but there are other upstream and downstream elements that do not overlap. There may also be a link between cytochrome c_{6A} and Ca^{2+} signalling, as HLA3 and LCIA have been shown to be in part regulated through calcium flux (Mackinder, 2018). LCIA showed a strong difference in regulation between cw15 and Uex55 upon high light exposure, but HLA3 showed a much weaker difference, therefore again cytochrome c_{6A} signalling could be interacting with the calcium retrograde signals rather than being the sole upstream messenger.

6.3.1.2. Cytochrome c_{6A} regulates motility under high light stress.

Motility demonstrated a strong downregulation upon HL exposure in cw15 that was diminished in Ue55. Analysis of the transcripts that were downregulated in this way also showed that many of the genes involved coded for structural components of the flagella and intraflagellar transport. This suggests that HL exposure led to a potential downregulation in flagellar subunits and repair, which was not observed without cytochrome c_{6A}. From this RNAseq data and current annotation of *C. reinhardtii* motility genes, it is difficult to determine from this data whether positive or negative phototaxis, or phototaxis overall, is being downregulated, and this warrants further assays of motility under HL exposure in Uex55 compared to cw15. One suggestion would be that positive phototaxis is being downregulated in response to high light, as the organism would prefer to move away from the source of stress until a more suitable light intensity is reached. As this downregulation did not occur when cytochrome c_{6A} was knocked out, it appears likely that this downregulation of positive phototaxis would be regulated by cytochrome c_{6A}, potentially as a result

of an over-reduced plastoquinone pool. The fact that the cultures were shaken throughout the experiment may also have affected phototaxis, as it meant that swimming in a certain direction would not lead to a change in light intensity experienced by the cell. It is also worth noting that as action potentials and Ca^{2+} flux are involved in phototactic regulation, there could be a link between CCM and motility through calcium signalling. However, photosynthesis has been shown to regulate motility (Takahashi and King, 2003), CCM and carotenoid synthesis, so cytochrome c_{6A} could be providing a signalling pathway not directly involving calcium.

6.3.1.3. Potential function for cytochrome c_{6A} in higher plants

Genes encoding cytochrome c_{6A} have been found in higher plants as well as eukaryotic green algae. However the methods of CCM in higher plants are very different from that of *C. reinhardtii* (Meyer *et al.*, 2016; Mackinder, 2018), and multicellular higher plants do not exhibit flagellar phototaxis. Therefore although the role of cytochrome c_{6A} in higher plants could be in retrograde signalling in response to HL, and one target of this signalling could be carotenoid biosynthesis, further study would need to look into any other responses under cytochrome c_{6A} control. In *Arabidopsis thaliana*, a group of plastids with a unique structure and proteome called "sensory plastids" were discovered in vascular parenchyma (Beltrán *et al.*, 2018). These plastids are thought to be involved in detecting environmental stress and relaying stress signals to the nucleus and the rest of the plant, and cytochrome c_{6A} was found in the proteome unique to these plastids. Although there has been no observation of sensory plastids in *C. reinhardtii*, and algal homologues of the other genes identified in these plastids by Beltrán *et al* (2018) showed no significant transcript regulation in this study (data not shown), cytochrome c_{6A} could play a specific signalling role in these plastids in higher plants.

6.3.2. RNA sequencing data does not allow analysis of the proteome

A limitation of this study is that it only analyses changes in the transcriptome of *C. reinhardtii* upon exposure to HL. Therefore any post-transcriptional, translational and post-translational modulation will not be observed. Many gene expression changes in *C. reinhardtii* do occur at the post-transcriptional level, particularly as a result of retrograde signalling, an example of which is the upregulation of light harvesting complexes upon HL stress (Floris *et al.*, 2013). This post-transcriptional control of the LHCs combined with the observation that the increase in LHC transcripts under HL conditions in *C. reinhardtii* are transient and return to a lower level by 6-8 hours (Durnford *et al.*, 2003) may explain why LHC transcripts were not highlighted in this study as

significantly differently regulated in cw15 and uex55, although there is also the possibility that cytochrome c_{6A} does not regulate the LHC. Therefore to gain a more holistic view of the effect of cytochrome c_{6A} on HL stress the proteome of the cultures will need to be analysed, for example through mass-spectrometry and peptide mass fingerprinting. A metabolomic analysis would also provide more insight into how anabolic and catabolic metabolism are affected under HL stress, and whether cytochrome c_{6A} plays a role in regulating those pathways.

6.3.3. Future directions

The RNA sequencing data generated in this study opens up many opportunities for future experiments regarding cytochrome c_{6A} function. Similar studies into the transcriptome of cytochrome c_{6A} knockout lines of higher plants upon HL stress should be performed, as well as other systems biology approaches to uex55 to analyse proteomic and metabolomic changes as a result of cytochrome c_{6A} signalling. For example, as carotenoid biosynthesis appears to be upregulated more in cw15 than in uex55 upon HL exposure, a pigment extraction and analysis could confirm cytochrome c_{6A} involvement in this pathway and by extension in NPQ. Quantitative reverse-transcriptase PCR could also confirm the transcript changes observed in this study and allow analysis of transcript level changes over multiple time points upon exposure. As motility appears to be regulated differently in cw15 and uex55 under HL stress, phototactic assays analysing the cultures' ability to exhibit positive and negative phototaxis could confirm this potential link. Equally, performing an assay to investigate the carbon concentrating efficiency of uex55 under HL conditions will provide more insight into what aspects of CCM cytochrome c_{6A} is involved in activating.

Overall this study has provided strong evidence that cytochrome c_{6A} is involved in retrograde signalling and transcriptional control of certain HL response genes. The key ontologies discovered thus far suggest that cytochrome c_{6A} upregulates carotenoid biosynthesis and CCM under HL, and downregulates motility, and therefore plays a key role in signalling within the photoprotective response.

7. Final conclusions and future work

7.1. Introduction

The results of this study have provided insight into the ancestry and function of cytochrome c_{6A} and its homologues cytochrome c_{6B} and c_{6C}. Phylogenetic analysis has provided evidence to show that cytochromes c_{6B} and c_{6C} are likely to be orthologues and therefore functionally identical, and I propose that the previous proteins identified as cytochrome c_{6C} should thus be named cytochrome c_{6B}. This abolishes the distinction between cytochromes c_{6B} and c_{6C}, but retains the separation from the eukaryotic homologues. This knowledge combined with extensive searches of photosynthetic organism ESTs and proteomes for cytochrome c_{6A} and c_{6B} homologues has allowed for an updated model of cytochrome c_6 family ancestry and evolution, which implies that cytochrome c_{6B} formed first through gene duplication of c_6 followed by a mutation to bring about the change in midpoint potential, and the LIP was subsequently inserted into the cytochrome c_{6B} coding region in the ancestor to the higher plant and green algae lineages to form cytochrome c_{6A}.

Characterisation of *Chlamydomonas reinhardtii* mutants with cytochrome c_{6A} either knocked out (Uex55) or potentially knocked down (Uex78 and Uex87) allowed for functional analysis of cytochrome c_{6A}. A possible disadvantage in growth under fluctuating light and the dark to light transition in the Uex mutant lines suggested a role in photoprotection for cytochrome c_{6A}, particularly in non-photochemical quenching as NPQ is fast enough to deal with fast changes in light conditions, and growth under constant high light (400 µmol $m^{-2}s^{-1}$) showed no clear phenotype. Chlorophyll fluorescence also showed a decrease in NPQ in the Uex mutants, particularly Uex55, compared to the background strain cw15, despite there being no noticeable change in photochemical quenching. Consistent with this, no major change in structure of cytochrome c_{6A} was observed using circular dichroism when the protein was exposed to a pH range of 2-7, indicating that it was not denatured and could potentially still function when the luminal pH fluctuated, which is a key requirement for fast-acting NPQ. However, these observations do not exclude more subtle changes in protein structure on pH change. Finally, transcriptomic analysis of cw15 and Uex55 upon high light exposure suggests that cytochrome c_{6A} contributed to the regulation of transcript levels of a number of genes, because removing cytochrome c_{6A} diminished this transcriptional regulation. This would be consistent with a signalling role of cytochrome c_{6A}. However it could also be an indirect effect where cytochrome c_{6A} loss led to the found changes in transcript levels through bioenergetics stress. Genes found to be differentially regulated thus far include those involved in carbon concentration, motility and carotenoid biosynthesis. The latter group of genes provides further evidence of an interaction between cytochrome c_{6A} and NPQ. CCM could also be seen as a

photoprotective element as it potentially provides a sink for reducing equivalents by enhancing carbon dioxide availability. However, we cannot exclude the possibility that the changes seen are a response to a disturbance resulting from the loss of an electron transfer pathway involving cytochrome $c_{6A.}$ It is also possible that cytochrome c_{6A} has both a 'safety valve' electron transfer and a signalling role, with the signalling responding to the redox state of cytochrome $c_{6A.}$

7.2. Previous work on cytochrome c_{6A} function

The work done in this study can be used in conjunction with previous work on cytochrome c_{6A} to help infer its function. However, it must be noted that a majority of previous work had been performed on the higher plant *Arabidopsis thaliana*. Many of the gene ontologies identified in RNA sequencing to be regulated directly or indirectly by cytochrome c_{6A} are unique to *C. reinhardtii* rather than higher plants, such as motility and CCM. Therefore, as higher plants also have cytochrome c_{6A} further study into high light regulation will need to be performed before firm conclusions can be made.

One study noted that a knockdown of plastocyanin combined with a knockout of cytochrome c_{6A} in *A. thaliana* was lethal (Gupta *et al.*, 2002). Therefore, without cytochrome c_{6A} being expressed the plants could not survive under low plastocyanin concentrations. This observation could fit well into the safety valve hypothesis, as cytochrome c_{6A} could alleviate the impairment of electron flow from the plastoquinone pool to PSI when plastocyanin is low, and therefore allow a stunted but viable line of plants. (However, if the route from cytochrome c_{6A} to PSI were through plastocyanin, see below, one might expect the reduction in plastocyanin concentration to be epistatic to loss of cytochrome c_{6A} .) The results of Gupta *et al.* would also be consistent with cytochrome c_{6A} playing a signalling role in photoprotection and possibly NPQ, because a plant with diminished NPQ response would rely on photochemical quenching and increased electron flow to prevent excessive singlet oxygen production. Therefore, below a certain threshold of plastocyanin expression, which could lead to a viable plant when NPQ is optimal, the mutation would prove lethal.

Another prior observation was that cytochrome c_{6A} was able to interact and reduce plastocyanin and PSI (Molina-Heredia *et al.*, 2003; Marcaida *et al.*, 2006). This would suggest involvement with an alternative electron pathway and therefore support the safety valve hypothesis. However, the kinetics of electron transfer to plastocyanin and PSI were considerably slower than for cytochrome c_6. This, combined with the low native expression level of cytochrome c_{6A} (Nimmo, 2011) suggests that this electron transfer may not have a large effect on photochemical quenching in

photosynthesis. This is supported by the negligible change in cytochrome c_{6A} transcript levels under high light stress and the lack of change in photochemical quenching observed through chlorophyll fluorescence. The interaction experiments were also performed *in vitro*, and there is no direct evidence for this interaction occurring *in vivo*.

7.3. Current working hypothesis for cytochrome c_{6A} function

Based on the data from this study, the leading hypothesis of cytochrome c_{6A} function is that it is likely to be a signalling protein for photoprotection and response to high light stress (although if it can also reduce PC and PSI under high light conditions, it could function both as a signalling protein and a safety valve). The fact that cytochrome c_{6A} is of such low abundance makes the possibility that its sole role is in physiologically significant electron transfer role less likely. Under a signalling model, when the plastoquinone pool is over-reduced, cytochrome c_{6A} would oxidise plastoquinol, then activate a downstream retrograde signalling pathway to induce a photoprotective response, primarily through providing a sink for reducing equivalents and ATP through CCM, and alterations in motility and NPQ (figure 7.1). Alternatively, a decrease in pH could alter the midpoint potential, thiol redox chemistry or surface charge distribution of cytochrome c_{6A} and therefore activate a downstream signalling response through new interactions only accessible under high light. The observation of cytochrome c_{6A} expression in the signalling plastids in *A. thaliana* adds support to the signalling hypothesis (Beltrán *et al.*, 2018).

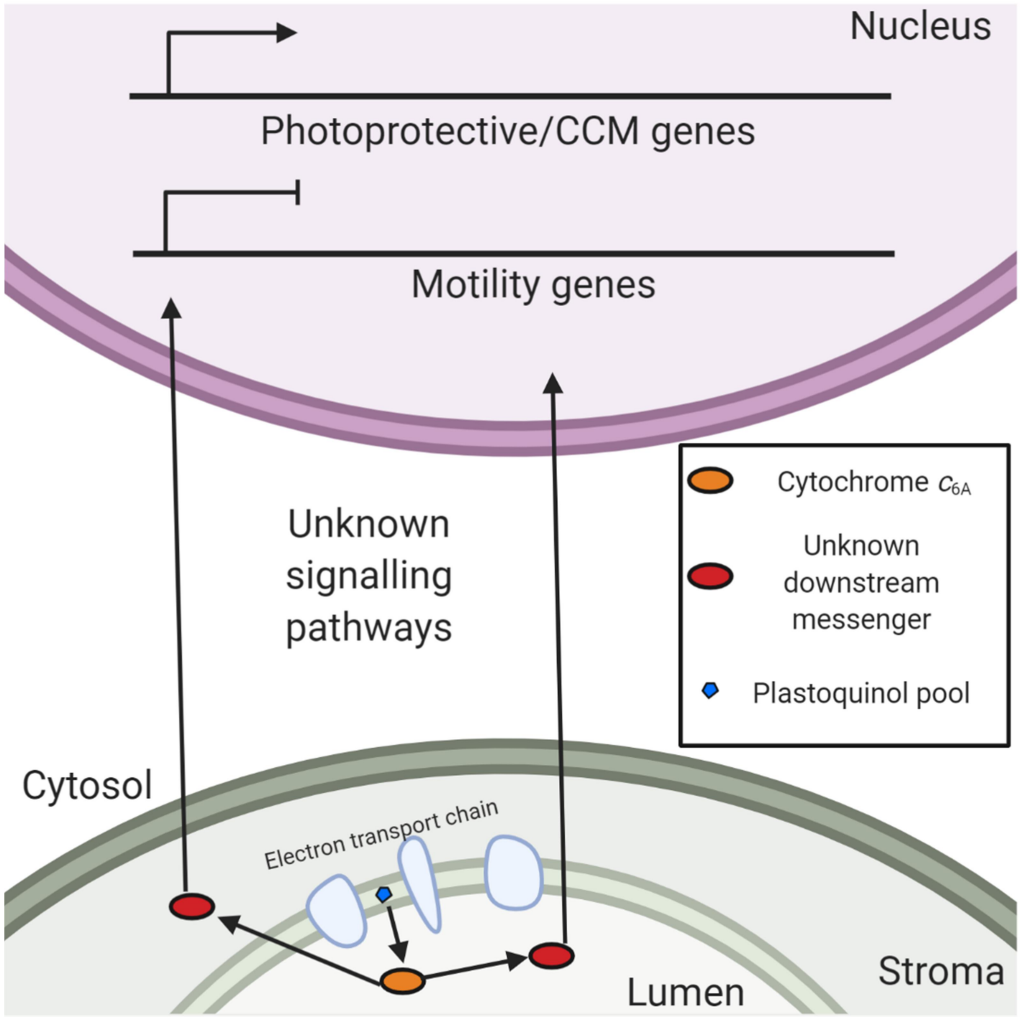


Figure 7.1: The current hypothesis for cytochrome c_{6A} signalling in C. reinhardtii. Under high light conditions an over-reduced plastoquinone pool reduces cytochrome c_{6A}, which can then activate a retrograde signalling pathway leading to transcriptional regulation.

The observation of a difference in growth rate between cw15 and the cytochrome c_{6A} mutant lines under heterotrophic but not photoautotrophic conditions could be linked to CCM. The presence of acetate in the medium appeared to cause a stronger trend of diminished growth in the cytochrome c_{6A} mutant lines under stress conditions (chapter 4). It was suggested that this could be because providing an external source of organic carbon resulted in lower levels of CO_2 fixation, and therefore without CO_2 fixation as a sink for reducing equivalents produced in the PETC more buildup of linear electron flow would occur leading to increased ROS production. In the RNA sequencing results it was observed that knocking out cytochrome c_{6A} led to the prevention of upregulation of CCM transcription (chapter 6). Loss of cytochrome c_{6A} would therefore interfere with the role of CO_2 fixation as an electron sink, and subsequently cause more stress in the cytochrome c_{6A} mutants in the presence of acetate.

The difference in mixotrophic and phototrophic growth of cytochrome c_{6A} mutant lines could also be explained if cytochrome c_{6A} is involved in state transition (qT) control. It has been shown that under mixotrophic conditions qT contributes more to photoprotection than qE (Finazzi *et al.*, 2006; Iwai *et al.*, 2007), therefore the diminished growth of Uex mutants under mixotrophic conditions could be explained if qT has been negatively affected in the mutant lines. It has also been shown that in *C. reinhardtii* qT contains greater shifts in LHCII than in higher plants, and happens over a shorter timescale (Vallon *et al.*, 1986; Allen, 1992). This means that the apparent drop in NPQ shown in the chlorophyll fluorescence studies could be attributed to a diminished qT, especially in the relaxation analysis that shows that NPQ recovery over the minutes timescale has been affected. Finally, genes involved in state transitions such as the STT7 kinase have been shown to have lower magnitudes of transcriptional change in the cytochrome c_{6A} knockout upon high light exposure. Therefore a potential direct or indirect role for cytochrome c_{6A} in signalling to state transition mechanisms should be investigated.

7.3.1. Potential challenges to the working hypothesis

Some experimental evidence does not directly fit a signalling role for cytochrome c_{6A}, although refining the model accordingly may help to resolve the conflict. For example, it was shown that cytochrome c_{6A} does not change conformation drastically upon reduction or oxidation (Marcaida *et al.*, 2006). This would rule out a signalling route whereby a major conformational change allowed reduced cytochrome c_{6A} to interact with a new partner (or dissociate from an existing one) to convey the signal further. However, cytochrome c_{6A} could instead pass this electron to a downstream molecule, which could then change conformation or be activated in some other way to induce the downstream response.

Another challenge to the hypothesis of a signalling role is that the growth curves of the cytochrome c_{6A} mutant lines in this study did not show a significant difference from the background strain, thus suggesting that under high light exposure, cytochrome c_{6A} did not provide a significant advantage to growth. This could be explained if CO_2 were not limited in the cultures and they were exposed to an even distribution of light. This would mean that diminished CCM or motility regulation in the knockout mutant lines would not be a disadvantage under the growth conditions.

The fact that NPQ is lowered but not eliminated in the absence of cytochrome c_{6A} suggested NPQ is not controlled by cytochrome c_{6A}. It should also be noted that a stronger difference in growth was seen between the Uex mutants and cw15 under fluctuating light and dark to light transition,

suggesting that cytochrome c_{6A} plays a shorter-term role in signalling in response to light changes. This would fit the chlorophyll fluorescence data and RNA sequencing results showing a diminished NPQ response in the knockout line Uex55. The RNA sequencing experiment was also performed over a timescale of 4 hours, which could be considered a short-term acclimation to high light relative to the multiple days of the growth curves, and in fact growth in the presence of Rose Bengal TAP showed a delayed adaptation to stress, but a similar eventual growth rate. A defective growth phenotype under fluctuating light conditions but not constant high light has been described in other studies, for example the study into terminal oxidase function in cyanobacteria (Lea-Smith *et al.*, 2013).

7.3.2. The NPQ phenotype of the Uex mutants happened faster than transcriptional regulation

The fact that NPQ was noticeably lower in the mutant lines suggests that the difference was not due to the upregulation of transcript levels, which works over a timescale of minutes to hours, as the chlorophyll fluorescence measurements would not have had enough time to regulate transcription in response to the actinic light turning on. However, the apparent disparity in kinetics could be explained if important NPQ genes such as *LHCSR.1/2* and *CVDE* had been upregulated in the wild type during growth of the culture before testing. On the other hand, the observed NPQ decrease without cytochrome c_{6A} could suggest that cytochrome c_{6A} is performing another role in more immediate NPQ (qE) in addition to the longer-term response of increasing carotenoid biosynthesis (qZ).

7.4. Future work

The results in this study suggest many new lines of experiments that could be performed to elucidate further the signalling role of cytochrome c_{6A}. First, the suggestion that cytochromes c_{6B} and c_{6C} are functionally the same needs to be tested outside of bioinformatics. Functional studies, for example a molecular genetics approach with knockout mutants for cytochrome c_{6B} in one of the species identified as containing a cytochrome c_{6B} gene by Bialek *et al.* (2008), should be performed and compared to those already performed on cytochrome c_{6C} (Vasudevan, 2019). As new databases and annotations of further photosynthetic organisms are published, searching for low redox midpoint cytochromes within the data will also continue to refine the current model of evolution of the cytochrome c_6 family. Further study into the dinoflagellate and proteobacterial potential cytochrome c_{6A} and c_{6B} proteins respectively will also confirm or disprove if they perform the same

function as cytochromes c_{6A} and c_{6B}, which could provide further insight into the overall function of these cytochromes.

The RNA sequencing data provides new paths of enquiry to follow regarding cytochrome c_{6A} function in photoprotection. Analysing the genes that appear most regulated by cytochrome c_{6A} for enrichment in certain promoter regions or regulatory sites associated with the genes would provide insight into potential downstream factors of cytochrome c_{6A} in the signalling pathway. Further analysis of the effect on expression of particularly interesting genes, such as CCM, carotenoid biosynthesis and flagellar proteins, should be confirmed under a wider range of conditions, for example through RT-qPCR. Finally, proteomic analysis of cytochrome c_{6A} knockout mutants under high light could provide insight into the regulation of genes that are not controlled at the transcriptional level. For example, the signalling cascade activated by cytochrome c_{6A} may affect not only transcript levels, but factors involved in protein synthesis and degradation, and post-translational modifications such as kinase activity.

If cytochrome c_{6A} does play a role in signalling, a key question will be what downstream messengers it interacts with. This would be difficult to measure *in vivo*, because interactions for electron will be too transient for detection through conventional methods such as yeast two hybrids. However, it may be possible to use cross-linking techniques to identify potential interaction partners, whose redox kinetics with cytochrome c_{6A} could be subsequently tested *in vitro* (Wetie *et al.*, 2014). The tests between these potential new interaction partners and cytochrome c_{6A} should also be performed under increasingly acidic pH, to determine if cytochrome c_{6A} alters binding in acidic conditions. Measurement of cytochrome c_{6A} redox midpoint potential under acidic pH should also be tested.

Further investigations should also be made to confirm cytochrome c_{6A} interaction with the processes highlighted by RNA sequencing. First, more investigation into NPQ could be performed using chlorophyll fluorescence. For example, measuring NPQ in cultures adapted to varying light intensities may show if exposure to high light does increase NPQ genes through cytochrome c_{6A}. Relaxation studies under a range of light intensities would also provide more insight into which element of NPQ cytochrome c_{6A} controls. Analysis of pigment content spectroscopically in the Uex mutants and cw15 under varying light intensity could also test a link between cytochrome c_{6A} and carotenoid biosynthesis. CCM interactions with cytochrome c_{6A} could be investigated by growing the Uex mutants and cw15 under varying CO_2 conditions. Oxygen evolution measurements could also be performed in a CO_2 depleted culture chamber, when the change in oxygen evolution upon incremental addition of bicarbonate would provide further evidence for or against cytochrome c_{6A}

involvement in CCM. In terms of motility, movement of cultures through low-percentage agar could be measured in the mutant lines compared to cytochrome c_{6A} both under standard and high light, in order to determine the effect of cytochrome c_{6A} on positive and negative phototaxis.

Transcriptomic and proteomic analysis should be performed on cytochrome c_{6A} knockout mutants under high light in higher plants, such as *A. thaliana*. It has been shown previously that knocking out cytochrome c_{6A} in *A. thaliana* led to a diminished growth under high light intensity (Vasudevan, 2019), therefore cytochrome c_{6A} could still be regulating photoprotective genes in response to high light intensities. However, as CCM and motility are different in *A. thaliana* than *C. reinhardtii*, the precise photoprotective genes may differ, which would be interesting to compare.

Finally, a construct exists that when transformed into *C. reinhardtii* will cause the overexpression of cytochrome c_{6A} (Slater, unpublished). The construct uses the native flanking regions of the *psaD* gene around the open reading frame of cytochrome c_{6A} to achieve this overexpression (Fischer and Rochaix, 2001). This construct could be used to see if overexpressing cytochrome c_{6A} has a phenotype in NPQ, CCM or motility that would support a signalling role for cytochrome c_{6A} in photoprotection. The construct can also be used in complementation experiments by transforming of Uex55 and testing if the mutant phenotypes have been reversed and wild type recovered.

Performing these experiments would provide further evidence for the function and ancestry of cytochrome c_{6A} and its homologues.

Bibliography

Allen, J.F., 1992. Protein phosphorylation in regulation of photosynthesis. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1098*(3), pp.275-335.

Amarnath, K., Zaks, J., Park, S.D., Niyogi, K.K. and Fleming, G.R., 2012. Fluorescence lifetime snapshots reveal two rapidly reversible mechanisms of photoprotection in live cells of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, *109*(22), pp.8405-8410.

An, F.Y. and Clewell, D.B., 1994. Characterization of the determinant (*traB*) encoding sex pheromone shutdown by the hemolysin/bacteriocin plasmid pAD1 in *Enterococcus faecalis*. *Plasmid*, *31*(2), pp.215-221.

Andersson, I. and Backlund, A., 2008. Structure and function of Rubisco. *Plant Physiology and Biochemistry*, 46(3), pp.275-291.

Anwaruzzaman, M., Chin, B.L., Li, X.P., Lohr, M., Martinez, D.A. and Niyogi, K.K., 2004. Genomic analysis of mutants affecting xanthophyll biosynthesis and regulation of photosynthetic light harvesting in *Chlamydomonas reinhardtii*. *Photosynthesis Research*, *82*(3), pp.265-276.

Aro, E.M., Virgin, I. and Andersson, B., 1993. Photoinhibition of photosystem II. Inactivation, protein damage and turnover. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1143*(2), pp.113-134.

Arraj, J.A. and Marinus, M.G., 1983. Phenotypic reversal in *dam* mutants of *Escherichia coli* K-12 by a recombinant plasmid containing the *dam+* gene. *Journal of bacteriology*, *153*(1), pp.562-565.

Arrieta, J., Barreira, A., Chioccioli, M., Polin, M. and Tuval, I., 2017. Phototaxis beyond turning: persistent accumulation and response acclimation of the microalga *Chlamydomonas reinhardtii*. *Scientific reports*, 7(1), p.3447.

Arslan, E., Schulz, H., Zufferey, R., Künzler, P. and Thöny-Meyer, L., 1998. Overproduction of the *Bradyrhizobium japonicum c*-Type Cytochrome Subunits of the cbb_3 Oxidase in *Escherichia coli*. *Biochemical and biophysical research communications*, *251*(3), pp.744-747.

Asada, K., Kiso, K. and Yoshikawa, K., 1974. Univalent reduction of molecular oxygen by spinach chloroplasts on illumination. *Journal of Biological Chemistry*, *249*(7), pp.2175-2181.
Asada, K., 1992. Ascorbate peroxidase–a hydrogen peroxide-scavenging enzyme in plants. *Physiologia Plantarum*, *85*(2), pp.235-241.

Asada, K., 2000. The water–water cycle as alternative photon and electron sinks. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1402), pp.1419-1431.

Ballottari, M., Dall'Osto, L., Morosinotto, T. and Bassi, R., 2007. Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *Journal of Biological Chemistry*, *282*(12), pp.8947-8958.

Ballottari, M., Truong, T.B., De Re, E., Erickson, E., Stella, G.R., Fleming, G.R., Bassi, R. and Niyogi, K.K., 2016. Identification of pH-sensing sites in the light harvesting complex stress-related 3 protein essential for triggering non-photochemical quenching in Chlamydomonas reinhardtii. *Journal of Biological Chemistry*, *291*(14), pp.7334-7346.

Baradaran, R., Berrisford, J.M., Minhas, G.S. and Sazanov, L.A., 2013. Crystal structure of the entire respiratory complex I. *Nature*, *494*(7438), p.443.

Barker, P.D. and Ferguson, S.J., 1999. Still a puzzle: why is haem covalently attached in *c*-type cytochromes? *Structure*, *7*(12), pp.R281-R290.

Baroli, I., Do, A.D., Yamane, T. and Niyogi, K.K., 2003. Zeaxanthin accumulation in the absence of a functional xanthophyll cycle protects *Chlamydomonas reinhardtii* from photooxidative stress. *The Plant Cell*, *15*(4), pp.992-1008.

Belatik, A., Joly, D., Hotchandani, S. and Carpentier, R., 2013. Re-evaluation of the side effects of cytochrome b_6f inhibitor dibromothymoquinone on photosystem II excitation and electron transfer. *Photosynthesis research*, *117*(1-3), pp.489-496.

Beltrán, J., Wamboldt, Y., Sanchez, R., LaBrant, E.W., Kundariya, H., Virdi, K.S., Elowsky, C. and Mackenzie, S.A., 2018. Specialized plastids trigger tissue-specific signaling for systemic stress response in plants. *Plant physiology*, 178(2), pp.672-683.

Bennoun, P., 1982. Evidence for a respiratory chain in the chloroplast. *Proceedings of the National Academy of Sciences*, *79*(14), pp.4352-4356.

Berman, S.A., Wilson, N.F., Haas, N.A. and Lefebvre, P.A., 2003. A novel MAP kinase regulates flagellar length in *Chlamydomonas*. *Current Biology*, 13(13), pp.1145-1149.

Berthold, P., Tsunoda, S.P., Ernst, O.P., Mages, W., Gradmann, D. and Hegemann, P., 2008. Channelrhodopsin-1 initiates phototaxis and photophobic responses in *Chlamydomonas* by immediate light-induced depolarization. *The Plant Cell*, 20(6), pp.1665-1677.

Betterle, N., Ballottari, M., Zorzan, S., de Bianchi, S., Cazzaniga, S., Dall'Osto, L., Morosinotto, T. and Bassi, R., 2009. Light-induced dissociation of an antenna hetero-oligomer is needed for non-photochemical quenching induction. *Journal of Biological Chemistry*, *284*(22), pp.15255-15266.

Bialek, W., Nelson, M., Tamiola, K., Kallas, T. and Szczepaniak, A., 2008. Deeply branching c_6-like cytochromes of cyanobacteria. *Biochemistry*, *47*(20), pp.5515-5522.

Bialek, W., Krzywda, S., Zatwarnicki, P., Jaskolski, M., Kolesinski, P. and Szczepaniak, A., 2014. Insights into the relationship between the haem-binding pocket and the redox potential of c_6 cytochromes: four atomic resolution structures of c_6 and c_6-like proteins from *Synechococcus* sp. PCC 7002. *Acta Crystallographica Section D: Biological Crystallography*, *70*(11), pp.2823-2832.

Blaby, I.K., Blaby-Haas, C.E., Pérez-Pérez, M.E., Schmollinger, S., Fitz-Gibbon, S., Lemaire, S.D. and Merchant, S.S., 2015. Genome-wide analysis on *Chlamydomonas reinhardtii* reveals the impact of hydrogen peroxide on protein stress responses and overlap with other stress transcriptomes. *The Plant Journal*, 84(5), pp.974-988.

Böhm, M., Boness, D., Fantisch, E., Erhard, H., Frauenholz, J., Kowalzyk, Z., Marcinkowski, N., Kateriya, S., Hegemann, P. and Kreimer, G., 2019. Channelrhodopsin-1 Phosphorylation Changes with Phototactic Behavior and Responds to Physiological Stimuli in *Chlamydomonas*. *The Plant Cell*, 31(4), pp.886-910.

Bonente, G., Ballottari, M., Truong, T.B., Morosinotto, T., Ahn, T.K., Fleming, G.R., Niyogi, K.K. and Bassi, R., 2011. Analysis of LhcSR3, a protein essential for feedback de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS biology*, *9*(1), p.e1000577.

Bonente, G., Pippa, S., Castellano, S., Bassi, R. and Ballottari, M., 2012. Acclimation of *Chlamydomonas reinhardtii* to different growth irradiances. *Journal of Biological Chemistry*, *287*(8), pp.5833-5847.

Boscov, J.S. and Feinleib, M.E., 1979. Phototactic response of *Chlamydomonas* to flashes of light–II. Response of individual cells. *Photochemistry and Photobiology*, *30*(4), pp.499-505.

Bowler, C., Van Camp, W., Van Montagu, M., Inzé, D. and Asada, K., 1994. Superoxide dismutase in plants. *Critical Reviews in Plant Sciences*, *13*(3), pp.199-218.

Briantais, J.M., Vernotte, C., Picaud, M. and Krause, G.H., 1979. A quantitative study of the slow decline of chlorophyll a fluorescence in isolated chloroplasts. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *548*(1), pp.128-138.

Brzezowski, P., Schlicke, H., Richter, A., Dent, R.M., Niyogi, K.K. and Grimm, B., 2014. The GUN4 protein plays a regulatory role in tetrapyrrole biosynthesis and chloroplast-to-nucleus signalling in *Chlamydomonas reinhardtii*. *The Plant Journal*, 79(2), pp.285-298.

Bukhov, N.G., Sridharan, G., Egorova, E.A. and Carpentier, R., 2003. Interaction of exogenous quinones with membranes of higher plant chloroplasts: modulation of quinone capacities as photochemical and non-photochemical quenchers of energy in Photosystem II during light–dark transitions. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1604*(2), pp.115-123.

Cardol, P., Forti, G. and Finazzi, G., 2011. Regulation of electron transport in microalgae. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1807*(8), pp.912-918.

Chain, R.K. and Malkin, R., 1979. On the interaction of 2, 5-dibromo-3-methyl-6-isopropylbenzoquinone (DBMIB) with bound electron carriers in spinach chloroplasts. *Archives of biochemistry and biophysics*, *197*(1), pp.52-56.

Chan, C.X., Soares, M.B., Bonaldo, M.F., Wisecaver, J.H., Hackett, J.D., Anderson, D.M., Erdner, D.L. and Bhattacharya, D., (2012). Analysis of *Alexandrium tamarense* (dinophyceae) genes reveals the complex evolutionary history of a microbial eukaryote. *Journal of phycology,* 48(5), pp.1130-1142. doi: 10.1111/j.1529-8817.2012.01194.

Chaux, F., Burlacot, A., Mekhalfi, M., Auroy, P., Blangy, S., Richaud, P. and Peltier, G., 2017. Flavodiiron proteins promote fast and transient O_2 photoreduction in *Chlamydomonas*. *Plant physiology*, *174*(3), pp.1825-1836.

Chen, Y.H., Yang, J.T. and Martinez, H.M., 1972. Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry*, *11*(22), pp.4120-4131.
Cheng, Z., Sattler, S., Maeda, H., Sakuragi, Y., Bryant, D.A. and DellaPenna, D., 2003. Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. *The Plant Cell*, 15(10), pp.2343-2356.

Chojnacki, S., Cowley, A., Lee, J., Foix, A. and Lopez, R., 2017. Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic acids research*, 45(W1), pp.W550-W553.
Cole, D.G., 2003. The intraflagellar transport machinery of *Chlamydomonas reinhardtii*. *Traffic*, 4(7), pp.435-442.

Cournac, L., Josse, E.M., Joët, T., Rumeau, D., Redding, K., Kuntz, M. and Peltier, G., 2000. Flexibility in photosynthetic electron transport: a newly identified chloroplast oxidase involved in chlororespiration. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1402), pp.1447-1454.

Couso, I., Vila, M., Vigara, J., Cordero, B.F., Vargas, M.Á., Rodríguez, H. and León, R., 2012. Synthesis of carotenoids and regulation of the carotenoid biosynthesis pathway in response to high light stress in the unicellular microalga *Chlamydomonas reinhardtii*. *European journal of phycology*, *47*(3), pp.223-232.

Cross, F.R., 2016. Tying down loose ends in the *Chlamydomonas* genome: functional significance of abundant upstream open reading frames. *G3: Genes, Genomes, Genetics*, *6*(2), pp.435-446.

Dang, K.V., Plet, J., Tolleter, D., Jokel, M., Cuiné, S., Carrier, P., Auroy, P., Richaud, P., Johnson, X., Alric, J. and Allahverdiyeva, Y., 2014. Combined increases in mitochondrial cooperation and oxygen photoreduction compensate for deficiency in cyclic electron flow in *Chlamydomonas reinhardtii*. *The Plant Cell*, *26*(7), pp.3036-3050.

Davies, D.R. and Plaskitt, A., 1971. Genetical and structural analyses of cell-wall formation in *Chlamydomonas reinhardi*. *Genetics Research*, *17*(1), pp.33-43.

Davis, M.C., Fiehn, O. and Durnford, D.G., 2013. Metabolic acclimation to excess light intensity in *Chlamydomonas reinhardtii*. *Plant, cell & environment*, *36*(7), pp.1391-1405.

Deleage, G. and Geourjon, C., 1993. An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Bioinformatics*, *9*(2), pp.197-199.

Depège, N., Bellafiore, S. and Rochaix, J.D., 2003. Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. *Science*, *299*(5612), pp.1572-1575.

Domozych, D., Ciancia, M., Fangel, J.U., Mikkelsen, M.D., Ulvskov, P. and Willats, W.G., 2012. The cell walls of green algae: a journey through evolution and diversity. *Frontiers in plant science*, *3*, p.82.

Dorrell, R.G. and Howe, C.J., 2015. Integration of plastids with their hosts: Lessons learned from dinoflagellates. *Proceedings of the National Academy of Sciences*, 112(33), pp.10247-10254.

Dumas, L., Zito, F., Blangy, S., Auroy, P., Johnson, X., Peltier, G. and Alric, J., 2017. A stromal region of cytochrome b_6f subunit IV is involved in the activation of the Stt7 kinase in *Chlamydomonas*. *Proceedings of the National Academy of Sciences*, 114(45), pp.12063-12068.

Duncan, O., van der Merwe, M.J., Daley, D.O. and Whelan, J., 2013. The outer mitochondrial membrane in higher plants. *Trends in plant science*, *18*(4), pp.207-217.

Durnford, D.G., Price, J.A., McKim, S.M. and Sarchfield, M.L., 2003. Light-harvesting complex gene expression is controlled by both transcriptional and post-transcriptional mechanisms during photoacclimation in *Chlamydomonas reinhardtii*. *Physiologia Plantarum*, *118*(2), pp.193-205.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), pp.1792-1797.

Elrad, D. and Grossman, A.R., 2004. A genome's-eye view of the light-harvesting polypeptides of *Chlamydomonas reinhardtii*. *Current genetics*, *45*(2), pp.61-75.

Endo, T. and Asada, K., 1996. Dark induction of the non-photochemical quenching of chlorophyll fluorescence by acetate in *Chlamydomonas reinhardtii*. *Plant and cell physiology*, *37*(4), pp.551-555.

Erickson, E., Wakao, S. and Niyogi, K.K., 2015. Light stress and photoprotection in *Chlamydomonas reinhardtii*. *The Plant Journal*, *82*(3), pp.449-465.

Fang, W., Si, Y., Douglass, S., Casero, D., Merchant, S.S., Pellegrini, M., Ladunga, I., Liu, P. and Spalding, M.H., 2012. Transcriptome-wide changes in *Chlamydomonas reinhardtii* gene expression regulated by carbon dioxide and the CO_2-concentrating mechanism regulator CIA5/CCM1. The Plant Cell, 24(5), pp.1876-1893.

Farmer, E.E. and Davoine, C., 2007. Reactive electrophile species. *Current opinion in plant biology*, *10*(4), pp.380-386.

Feenstra, R.P. and Tseng, S.C., 1992. What is actually stained by Rose Bengal?. *Archives of ophthalmology*, *110*(7), pp.984-993.

Felsenstein, J., 1989. PHYLIP (Version 3.6) Phylogeny Inference Package. *Cladistics*, 5, pp.164-166.

Ferenczi, A., Pyott, D.E., Xipnitou, A. and Molnar, A., 2017. Efficient targeted DNA editing and replacement in *Chlamydomonas reinhardtii* using Cpf1 ribonucleoproteins and single-stranded DNA. *Proceedings of the National Academy of Sciences*, *114*(51), pp.13567-13572.

Fett, J.P. and Coleman, J.R., 1994. Regulation of periplasmic carbonic anhydrase expression in *Chlamydomonas reinhardtii* by acetate and pH. *Plant physiology*, *106*(1), pp.103-108.

Finazzi, G., Furia, A., Barbagallo, R.P. and Forti, G., 1999. State transitions, cyclic and linear electron transport and photophosphorylation in *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1413*(3), pp.117-129.

Finazzi, G., Johnson, G.N., Dall'Osto, L., Zito, F., Bonente, G., Bassi, R. and Wollman, F.A., 2006. Nonphotochemical quenching of chlorophyll fluorescence in *Chlamydomonas reinhardtii. Biochemistry*, *45*(5), pp.1490-1498.

Fischer, N. and Rochaix, J.D., 2001. The flanking regions of *PsaD* drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Molecular Genetics and Genomics*, *265*(5), pp.888-894.

Fischer, B.B., Krieger-Liszkay, A. and Eggen, R.I., 2004. Photosensitizers neutral red (type I) and Rose Bengal (type II) cause light-dependent toxicity in *Chlamydomonas reinhardtii* and induce the *Gpxh* gene via increased singlet oxygen formation. *Environmental science & technology*, *38*(23), pp.6307-6313.

Fischer, B.B., Krieger-Liszkay, A. and Eggen, R.I., 2005. Oxidative stress induced by the photosensitizers neutral red (type I) or Rose Bengal (type II) in the light causes different molecular responses in *Chlamydomonas reinhardtii*. *Plant science*, *168*(3), pp.747-759.

Fischer, B.B., Eggen, R.I., Trebst, A. and Krieger-Liszkay, A., 2006. The glutathione peroxidase homologous gene *Gpxh* in *Chlamydomonas reinhardtii* is upregulated by singlet oxygen produced in photosystem II. *Planta*, *223*(3), pp.583-590.

Fischer, B.B., Krieger-Liszkay, A., Hideg, É., Šnyrychová, I., Wiesendanger, M. and Eggen, R.I., 2007. Role of singlet oxygen in chloroplast to nucleus retrograde signaling in *Chlamydomonas reinhardtii*. *FEBS letters*, 581(29), pp.5555-5560.

Fischer, B.B., Dayer, R., Schwarzenbach, Y., Lemaire, S.D., Behra, R., Liedtke, A. and Eggen, R.I., 2009. Function and regulation of the glutathione peroxidase homologous gene *GPXH/GPX5* in *Chlamydomonas reinhardtii*. *Plant molecular biology*, *71*(6), pp.569-583.

Fischer, B.B., Ledford, H.K., Wakao, S., Huang, S.G., Casero, D., Pellegrini, M., Merchant, S.S., Koller, A., Eggen, R.I. and Niyogi, K.K., 2012. SINGLET OXYGEN RESISTANT 1 links reactive electrophile signaling to singlet oxygen acclimation in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, *109*(20), pp.E1302-E1311.

Floris, M., Bassi, R., Robaglia, C., Alboresi, A. and Lanet, E., 2013. Post-transcriptional control of light-harvesting genes expression under light stress. *Plant molecular biology*, 82(1-2), pp.147-154.

Forti, G., Furia, A., Bombelli, P. and Finazzi, G., 2003. In vivo changes of the oxidation-reduction state of NADP and of the ATP/ADP cellular ratio linked to the photosynthetic activity in *Chlamydomonas reinhardtii*. Plant physiology, 132(3), pp.1464-1474.

Garrity, G., Staley, J.T., Boone, D.R., De Vos, P., Goodfellow, M., Rainey, F.A., Garrity, G.M. and Schleifer, K.H., 2006. *Bergey's Manual® of systematic bacteriology: volume two: the proteobacteria*. Springer Science & Business Media.

Giorgi, F.M., Del Fabbro, C. and Licausi, F., 2013. Comparative study of RNA-seq-and microarray-derived coexpression networks in *Arabidopsis thaliana*. *Bioinformatics*, *29*(6), pp.717-724.

Goodenough, U.W., 1970. Chloroplast division and pyrenoid formation in *Chlamydomonas reinhardtii*. *Journal of Phycology*, *6*(1), pp.1-6.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S., 2011. Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, *40*(D1), pp.D1178-D1186.

Gorelova, V., Bastien, O., De Clerck, O., Lespinats, S., Rébeillé, F. and Van Der Straeten, D., 2019. Evolution of folate biosynthesis and metabolism across algae and land plant lineages. *Scientific reports*, 9(1), p.5731.

Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols*, *1*(6), p.2876.

Gupta, R., He, Z. and Luan, S., 2002. Functional relationship of cytochrome c_6 and plastocyanin in *Arabidopsis*. *Nature*, *417*(6888), p.567.

Hackett, J.D., Scheetz, T.E., Yoon, H.S., Soares, M.B., Bonaldo, M.F., Casavant, T.L. and Bhattacharya, D., (2005). Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC genomics*, *6*(1), p.80. doi: 10.1186/1471-2164-6-80.

Hall, M., Mata-Cabana, A., Åkerlund, H.E., Florencio, F.J., Schröder, W.P., Lindahl, M. and Kieselbach, T., 2010. Thioredoxin targets of the plant chloroplast lumen and their implications for plastid function. *Proteomics*, *10*(5), pp.987-1001.

Hall, B. G. (2017) Phylogenetic trees made easy: A how to manual. Sinauer, 5th edition.

Halliwell, B., 1984. *Chloroplast metabolism: the structure and function of chloroplasts in green leaf cells.* (No. 04; QK882, H3 1984.) pp 180-202.

Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology*, *166*(4), pp.557-580.

Havaux, M., Eymery, F., Porfirova, S., Rey, P. and Dörmann, P., 2005. Vitamin E protects against photoinhibition and photooxidative stress in *Arabidopsis thaliana*. *The Plant Cell*, *17*(12), pp.3451-3469.

Hertle, A.P., Blunder, T., Wunder, T., Pesaresi, P., Pribil, M., Armbruster, U. and Leister, D., 2013. PGRL1 is the elusive ferredoxin-plastoquinone reductase in photosynthetic cyclic electron flow. *Molecular cell*, *49*(3), pp.511-523.

Hieber, A.D., Bugos, R.C. and Yamamoto, H.Y., 2000. Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1482*(1-2), pp.84-91.

Holland, E.M., Harz, H., Uhl, R. and Hegemann, P., 1997. Control of phobic behavioral responses by rhodopsin-induced photocurrents in *Chlamydomonas*. *Biophysical journal*, 73(3), pp.1395-1401.

Holub, O., Seufferheld, M.J., Gohlke, C., Govindjee, Heiss, G.J. and Clegg, R.M., 2007. Fluorescence lifetime imaging microscopy of *Chlamydomonas reinhardtii*: non-photochemical quenching mutants and the effect of photosynthetic inhibitors on the slow chlorophyll fluorescence transient. *Journal of Microscopy*, *226*(2), pp.90-120.

Houille-Vernes, L., Rappaport, F., Wollman, F.A., Alric, J. and Johnson, X., 2011. Plastid terminal oxidase 2 (PTOX2) is the major oxidase involved in chlororespiration in *Chlamydomonas*. *Proceedings of the National Academy of Sciences*, *108*(51), pp.20820-20825.

Howe, C.J., Schlarb-Ridley, B.G., Wastl, J., Purton, S. and Bendall, D.S., 2005. The novel cytochrome c_6 of chloroplasts: a case of evolutionary bricolage? *Journal of experimental botany*, *57*(1), pp.13-22.
Howe, C.J., Nimmo, R.H., Barbrook, A.C. and Bendall, D.S., 2016. Cytochrome c_{6A} of chloroplasts. In *Cytochrome Complexes: Evolution, Structures, Energy Transduction, and Signaling* (pp. 701-712). Springer, Dordrecht.

Im, C.S. and Grossman, A.R., 2002. Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. *The Plant Journal*, *30*(3), pp.301-313.

Iwai, M., Kato, N. and Minagawa, J., 2007. Distinct physiological responses to a high light and low CO_2 environment revealed by fluorescence quenching in photoautotrophically grown *Chlamydomonas reinhardtii*. *Photosynthesis research*, *94*(2-3), pp.307-314.

Jacobs, J., Pudollek, S., Hemschemeier, A. and Happe, T., 2009. A novel, anaerobically induced ferredoxin in *Chlamydomonas reinhardtii.* FEBS letters, 583(2), pp.325-329.

Jagendorf, A.T. and Uribe, E., 1966. ATP formation caused by acid-base transition of spinach chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*, *55*(1), p.170.

Jans, F., Mignolet, E., Houyoux, P.A., Cardol, P., Ghysels, B., Cuiné, S., Cournac, L., Peltier, G., Remacle, C. and Franck, F., 2008. A type II NAD (P) H dehydrogenase mediates light-independent plastoquinone reduction in the chloroplast of *Chlamydomonas*. *Proceedings of the National Academy of Sciences*, *105*(51), pp.20546-20551.

Janssen, M., Janssen, M., de Winter, M., Tramper, J., Mur, L.R., Snel, J. and Wijffels, R.H., 2000. Efficiency of light utilization of *Chlamydomonas reinhardtii* under medium-duration light/dark cycles. *Journal of Biotechnology*, *78*(2), pp.123-137.

Jiang, W., Brueggeman, A.J., Horken, K.M., Plucinak, T.M. and Weeks, D.P., 2014. Successful transient expression of Cas9 and single guide RNA genes in *Chlamydomonas reinhardtii*. *Eukaryotic cell*, *13*(11), pp.1465-1469.

Johnson, X. and Alric, J., 2012. Interaction between starch breakdown, acetate assimilation, and photosynthetic cyclic electron flow in *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*, *287*(31), pp.26445-26452.

Johnson, X., Steinbeck, J., Dent, R.M., Takahashi, H., Richaud, P., Ozawa, S.I., Houille-Vernes, L., Petroutsos, D., Rappaport, F., Grossman, A.R. and Niyogi, K.K., 2014. Proton gradient regulation 5-mediated cyclic electron flow under ATP-or redox-limited conditions: a study of *ΔATpase pgr5* and *ΔrbcL pgr5* mutants in the green alga *Chlamydomonas reinhardtii*. *Plant physiology*, *165*(1), pp.438-452.

Jones, D.T., Taylor, W.R. and Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*, *8*(3), pp.275-282.

Jordan, D.B. and Ogren, W.L., 1984. The CO 2/O 2 specificity of ribulose 1, 5-bisphosphate carboxylase/oxygenase. *Planta*, 161(4), pp.308-313.

Josef, K., Saranak, J. and Foster, K.W., 2006. Linear systems analysis of the ciliary steering behavior associated with negative-phototaxis in *Chlamydomonas reinhardtii*. *Cell motility and the cytoskeleton.* 63(12), pp.758-777.

Kallenbach, N.R., Lyu, P. and Zhou, H., 1996. CD spectroscopy and the helix-coil transition in peptides and polypeptides. In Circular dichroism and the conformational analysis of biomolecules (pp. 201-259). Springer, Boston, MA.

Kautsky, H., Appel, W. and Amann, H., 1960. Chlorophyll fluorescence and carbon assimilation. Part XIII. The fluorescence and the photochemistry of plants. *Biochemische zeitschrift*, *332*, pp.277-292.

Kawai, Y., Ono, E. and Mizutani, M., 2014. Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, 78(2), pp.328-343.

Keeling, P.J., 2010. The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1541), pp.729-748.

Kelly, S.M., Jess, T.J. and Price, N.C., 2005. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1751*(2), pp.119-139.

Kerfeld, C.A. and Krogmann, D.W., 1998. Photosynthetic cytochromes *c* in cyanobacteria, algae, and plants. *Annual review of plant biology*, *49*(1), pp.397-425.

Kim, B.H., Kim, S., Kim, H.G., Lee, J., Lee, I.S. and Park, Y.K., 2005. The formation of cyclopropane fatty acids in *Salmonella enterica serovar Typhimurium*. Microbiology, 151(1), pp.209-218.

Kluth, J., Schmidt, A., März, M., Krupinska, K. and Lorbiecke, R., 2012. *Arabidopsis* Zinc Ribbon 3 is the ortholog of yeast mitochondrial HSP70 escort protein HEP1 and belongs to an ancient protein family in mitochondria and plastids. *FEBS letters*, 586(19), pp.3071-3076.

Koch, K., 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current opinion in plant biology*, 7(3), pp.235-246.

Kosourov, S., Jokel, M., Aro, E.M. and Allahverdiyeva, Y., 2018. A new approach for sustained and efficient H_2 photoproduction by *Chlamydomonas reinhardtii*. *Energy & Environmental Science*, *11*(6), pp.1431-1436.

Kramer, D.M., Sacksteder, C.A. and Cruz, J.A., 1999. How acidic is the lumen?. *Photosynthesis research*, *60*(2-3), pp.151-163.

Krause, G.H. and Weis, E., 1991. Chlorophyll fluorescence and photosynthesis: the basics. *Annual review of plant biology*, *42*(1), pp.313-349.

Krieger, A. and Weis, E., 1993. The role of calcium in the pH-dependent control of photosystem II. *Photosynthesis Research*, *37*(2), pp.117-130.

Krieger-Liszkay, A., 2005. Singlet oxygen production in photosynthesis. *Journal of experimental botany*, *56*(411), pp.337-346.

Kuchitsu, K., Tsuzuki, M. and Miyachi, S., 1988. Characterization of the pyrenoid isolated from unicellular green alga *Chlamydomonas reinhardtii*: particulate form of RuBisCO protein. *Protoplasma*, 144(1), pp.17-24.

Kühne, H., Szalai, V.A. and Brudvig, G.W., 1999. Competitive binding of acetate and chloride in photosystem II. *Biochemistry*, *38*(20), pp.6604-6613.

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, *33*(7), pp.1870-1874.

Kusai, A. and Yamanaka, T., 1973. Cytochrome *c* (553, *Chlorobium thiosulfatophilum*) is a sulphide-cytochrome *c* reductase. *FEBS letters*, *34*(2), pp.235-237.

Lalibertè, G. and de la Noüie, J., 1993. Auto-, Hetero-, And Mixotrophic Growth Of *Chlamydomonas Humicola* (Cmloroimiyckak) On Acetate 1. *Journal of phycology*, *29*(5), pp.612-620.

Lea-Smith, D.J., Ross, N., Zori, M., Bendall, D.S., Dennis, J.S., Scott, S.A., Smith, A.G. and Howe, C.J., 2013. Thylakoid terminal oxidases are essential for the cyanobacterium *Synechocystis* sp. PCC 6803 to survive rapidly changing light intensities. *Plant physiology*, *162*(1), pp.484-495.

Ledford, H.K., Chin, B.L. and Niyogi, K.K., 2007. Acclimation to singlet oxygen stress in *Chlamydomonas reinhardtii*. *Eukaryotic cell*, *6*(6), pp.919-930.

Lemeille, S., Willig, A., Depège-Fargeix, N., Delessert, C., Bassi, R. and Rochaix, J.D., 2009. Analysis of the chloroplast protein kinase Stt7 during state transitions. *PLoS biology*, *7*(3), p.e1000045.

Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S. and Wang, X., 1997. Cytochrome *c* and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell*, *91*(4), pp.479-489.

Li, X.P., Gilmore, A.M., Caffarri, S., Bassi, R., Golan, T., Kramer, D. and Niyogi, K.K., 2004. Regulation of photosynthetic light harvesting involves intrathylakoid lumen pH sensing by the PsbS protein. *Journal of Biological Chemistry*, *279*(22), pp.22866-22874.

Li, Z., Keasling, J.D. and Niyogi, K.K., 2012. Overlapping photoprotective function of vitamin E and carotenoids in *Chlamydomonas*. *Plant physiology*, *158*(1), pp.313-323.

Li, Z., Peers, G., Dent, R.M., Bai, Y., Yang, S.Y., Apel, W., Leonelli, L. and Niyogi, K.K., 2016. Evolution of an atypical de-epoxidase for photoprotection in the green lineage. *Nature plants*, *2*(10), p.16140.

Liguori, N., Roy, L.M., Opacic, M., Durand, G. and Croce, R., 2013. Regulation of light harvesting in the green alga *Chlamydomonas reinhardtii*: the C-terminus of LHCSR is the knob of a dimmer switch. *Journal of the American Chemical Society*, *135*(49), pp.18339-18342.

Luo, M., Cao, M., Kan, Y., Li, G., Snell, W. and Pan, J., 2011. The phosphorylation state of an aurora-like kinase marks the length of growing flagella in *Chlamydomonas*. *Current Biology*, 21(7), pp.586-591.

Mackinder, L.C., 2018. The *Chlamydomonas* CO_2-concentrating mechanism and its potential for engineering photosynthesis in plants. *New Phytologist*, 217(1), pp.54-61.

Mäkinen, O.E., Zannini, E., Koehler, P. and Arendt, E.K., 2016. Heat-denaturation and aggregation of quinoa (*Chenopodium quinoa*) globulins as affected by the pH value. *Food chemistry*, *196*, pp.17-24.

Malnoë, A., Wang, F., Girard-Bascou, J., Wollman, F.A. and de Vitry, C., 2014. Thylakoid FtsH protease contributes to photosystem II and cytochrome b_6f remodeling in *Chlamydomonas reinhardtii* under stress conditions. *The Plant Cell*, *26*(1), pp.373-390.

Mantovani, R.A., Fattori, J., Michelon, M. and Cunha, R.L., 2016. Formation and pH-stability of whey protein fibrils in the presence of lecithin. *Food Hydrocolloids*, *60*, pp.288-298.

Marcaida, M.J., Schlarb-Ridley, B.G., Worrall, J.A., Wastl, J., Evans, T.J., Bendall, D.S., Luisi, B.F. and Howe, C.J., 2006. Structure of cytochrome c_{6A}, a novel dithio-cytochrome of *Arabidopsis thaliana*,

and its reactivity with plastocyanin: implications for function. *Journal of molecular biology*, *360*(5), pp.968-977.

Marcus, Y., Schuster, G., Michaels, A. and Kaplan, A., 1986. Adaptation to CO_2 level and changes in the phosphorylation of thylakoid proteins during the cell cycle of *Chlamydomonas reinhardtii*. *Plant physiology*, *80*(2), pp.604-607.

Maruyama, S., Tokutsu, R. and Minagawa, J., 2014. Transcriptional regulation of the stress-responsive light harvesting complex genes in *Chlamydomonas reinhardtii*. *Plant and Cell Physiology*, 55(7), pp.1304-1310.

Mason, J.M., Bendall, D.S., Howe, C.J. and Worrall, J.A., 2012. The role of a disulfide bridge in the stability and folding kinetics of *Arabidopsis thaliana* cytochrome c_{6A}. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1824*(2), pp.311-318.

Maxwell, K. and Johnson, G.N., 2000. Chlorophyll fluorescence—a practical guide. *Journal of experimental botany*, *51*(345), pp.659-668.

Mayer, A.M., 1968. *Chlamydomonas*: adaptation phenomena in phototaxis. *Nature*, 217(5131), p.875.

McKim, S.M. and Durnford, D.G., 2006. Translational regulation of light-harvesting complex expression during photoacclimation to high-light in *Chlamydomonas reinhardtii*. *Plant Physiology and Biochemistry*, *44*(11-12), pp.857-865.

Melkonian, M. and Robenek, H., 1980. Eyespot membranes of *Chlamydomonas reinhardii*: a freeze-fracture study. *Journal of ultrastructure research*, *72*(1), pp.90-102.

Mendez-Alvarez, S., Leisinger, U. and Eggen, R.I., 1999. Adaptive responses in *Chlamydomonas reinhardtii*. *International Microbiology*, 2(1), pp.15-22.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L. and Marshall, W.F., 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*, *318*(5848), pp.245-250.

Mettler, T., Mühlhaus, T., Hemme, D., Schöttler, M.A., Rupprecht, J., Idoine, A., Veyel, D., Pal, S.K., Yaneva-Roder, L., Winck, F.V. and Sommer, F., 2014. Systems analysis of the response of photosynthesis, metabolism, and growth to an increase in irradiance in the photosynthetic model organism *Chlamydomonas reinhardtii*. *The Plant Cell*, *26*(6), pp.2310-2350.

Meyer, M.T., McCormick, A.J. and Griffiths, H., 2016. Will an algal CO_2-concentrating mechanism work in higher plants?. *Current Opinion in Plant Biology*, 31, pp.181-188.

Minagawa, J., 2011. State transitions—the molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1807*(8), pp.897-905.

Mitchell, M.C., Meyer, M.T. and Griffiths, H., 2014. Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized *Chlamydomonas reinhardtii*. *Plant Physiology*, *166*(2), pp.1073-1082.

Miura, K., Yamano, T., Yoshioka, S., Kohinata, T., Inoue, Y., Taniguchi, F., Asamizu, E., Nakamura, Y., Tabata, S., Yamato, K.T. and Ohyama, K., 2004. Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant physiology*, *135*(3), pp.1595-1607.

Molina-Heredia, F.P., Wastl, J., Navarro, J.A., Bendall, D.S., Hervás, M., Howe, C.J. and Miguel, A., 2003. Photosynthesis (communication arising): A new function for an old cytochrome? *Nature*, *424*(6944), p.33.

Moreland, D.E., 1980. Mechanisms of action of herbicides. *Annual review of plant physiology*, *31*(1), pp.597-638.

Morita, E., Kuroiwa, H., Kuroiwa, T. and Nozaki, H., 1997. High localization of ribulose-1, 5-bisphosphate carboxulase/oxygenase in the pyrenoids of *Chlamydomonas reinhardtii* (chlorophyta), as revealed by cryofixation and immunogold electron microscopy. *Journal of Phycology*, 33(1), pp.68-72.

Moroney, J.V., Husic, H.D. and Tolbert, N.E., 1985. Effect of carbonic anhydrase inhibitors on inorganic carbon accumulation by *Chlamydomonas reinhardtii*. *Plant Physiology*, 79(1), pp.177-183.

Mukherjee, A., Lau, C.S., Walker, C.E., Rai, A.K., Prejean, C.I., Yates, G., Emrich-Mills, T., Lemoine, S.G., Vinyard, D.J., Mackinder, L.C. and Moroney, J.V., 2019. Thylakoid localized bestrophin-like proteins are essential for the CO_2 concentrating mechanism of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, *116*(34), pp.16915-16920.

Müller, P., Li, X.P. and Niyogi, K.K., 2001. Non-photochemical quenching. A response to excess light energy. *Plant physiology*, *125*(4), pp.1558-1566.

Murata, N., Allakhverdiev, S.I. and Nishiyama, Y., 2012. The mechanism of photoinhibition in vivo: re-evaluation of the roles of catalase, α-tocopherol, non-photochemical quenching, and electron transport. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1817*(8), pp.1127-1133.

Mus, F., Cournac, L., Cardettini, V., Caruana, A. and Peltier, G., 2005. Inhibitor studies on non-photochemical plastoquinone reduction and H_2 photoproduction in *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1708*(3), pp.322-332.

Neckers, D.C., 1989. Rose Bengal. *Journal of Photochemistry and Photobiology A: Chemistry*, *47*(1), pp.1-29.

Ngounou Wetie, A.G., Sokolowska, I., Woods, A.G., Roy, U., Loo, J.A. and Darie, C.C., 2013. Investigation of stable and transient protein–protein interactions: past, present, and future. *Proteomics*, *13*(3-4), pp.538-557.

Nguyen, H.M., Baudet, M., Cuine, S., Adriano, J.M., Barthe, D., Billon, E., Bruley, C., Beisson, F., Peltier, G., Ferro, M. and Li-Beisson, Y., 2011. Proteomic profiling of oil bodies isolated from the unicellular green microalga *Chlamydomonas reinhardtii*: with focus on proteins involved in lipid metabolism. *Proteomics*, 11(21), pp.4266-4273.

Nilkens, M., Kress, E., Lambrev, P., Miloslavina, Y., Müller, M., Holzwarth, A.R. and Jahns, P., 2010. Identification of a slowly inducible zeaxanthin-dependent component of non-photochemical

quenching of chlorophyll fluorescence generated under steady-state conditions in *Arabidopsis*. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1797*(4), pp.466-475.

Nimmo, R.H., 2011. Location and function of cytochrome c_{6A}. PhD thesis, University of Cambridge.

Niyogi, K.K., Bjorkman, O. and Grossman, A.R., 1997a. *Chlamydomonas* xanthophyll cycle mutants identified by video imaging of chlorophyll fluorescence quenching. *The Plant Cell*, *9*(8), pp.1369-1380.

Niyogi, K.K., Bjorkman, O. and Grossman, A.R., 1997a. Chlamydomonas xanthophyll cycle mutants identified by video imaging of chlorophyll fluorescence quenching. *The Plant Cell*, *9*(8), pp.1369-1380.

Niyogi, K.K., Björkman, O. and Grossman, A.R., 1997b. The roles of specific xanthophylls in photoprotection. *Proceedings of the National Academy of Sciences*, *94*(25), pp.14162-14167.

Niyogi, K.K., Shih, C., Chow, W.S., Pogson, B.J., DellaPenna, D. and Björkman, O., 2001. Photoprotection in a zeaxanthin-and lutein-deficient double mutant of *Arabidopsis*. *Photosynthesis Research*, *67*(1-2), pp.139-145.

Nomura, C. T. (2001) *Electron Transport Proteins Of Synechococcus SP. PCC 7002*, *n/a*. Available at: https://etda.libraries.psu.edu/files/final_submissions/1251.

Ohnishi, N., Mukherjee, B., Tsujikawa, T., Yanase, M., Nakano, H., Moroney, J.V. and Fukuzawa, H., 2010. Expression of a Low CO_2–inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. *The Plant Cell*, 22(9), pp.3105-3117.

Pan, J., Wang, Q. and Snell, W.J., 2004. An aurora kinase is essential for flagellar disassembly in *Chlamydomonas*. *Developmental cell*, 6(3), pp.445-451.

Papageorgiou, G.C. ed., 2007. Chlorophyll a fluorescence: a signature of photosynthesis (Vol. 19). Springer Science & Business Media.

Papageorgiou, G.C., 2014. The non-photochemical quenching of the electronically excited state of chlorophyll a in plants: definitions, timelines, viewpoints, open questions. In *Non-Photochemical*

Quenching and Energy Dissipation in Plants, Algae and Cyanobacteria (pp. 1-44). Springer, Dordrecht.

Patterson, C.P. and Myers, J., 1973. Photosynthetic production of hydrogen peroxide by *Anacystis nidulans*. *Plant physiology*, 51(1), pp.104-109.

Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M. and Niyogi, K.K., 2009. An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature*, *462*(7272), p.518.

Pfannschmidt, T., 2003. Chloroplast redox signals: how photosynthesis controls its own genes. *Trends in plant science*, *8*(1), pp.33-41.

Pollock, S.V., Prout, D.L., Godfrey, A.C., Lemaire, S.D. and Moroney, J.V., 2004. The *Chlamydomonas reinhardtii* proteins Ccp1 and Ccp2 are required for long-term growth, but are not necessary for efficient photosynthesis, in a low-CO_2 environment. *Plant molecular biology*, 56(1), pp.125-132.

Polukhina, I., Fristedt, R., Dinc, E., Cardol, P. and Croce, R., 2016. Carbon supply and photoacclimation cross talk in the green alga *Chlamydomonas reinhardtii*. *Plant physiology*, *172*(3), pp.1494-1505.

Pribil, M., Pesaresi, P., Hertle, A., Barbato, R. and Leister, D., 2010. Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow. *PLoS biology*, *8*(1), p.e1000288.

Puthiyaveetil, S., Ibrahim, I.M. and Allen, J.F., 2013. Evolutionary rewiring: a modified prokaryotic gene-regulatory pathway in chloroplasts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1622), p.20120260.

Randall, J.T., 1969. The flagellar apparatus as a model organelle for the study of growth and morphopoiesis. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *173*(1030), pp.31-55.

Raven, J.A., Cockell, C.S. and De La Rocha, C.L., 2008. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1504), pp.2641-2650.

Retief, J.D., 2000. Phylogenetic analysis using PHYLIP. In *Bioinformatics methods and protocols* (pp. 243-258). Humana Press, Totowa, NJ.

Riaño-Pachón, D.M., Corrêa, L.G.G., Trejos-Espinosa, R. and Mueller-Roeber, B., 2008. Green transcription factors: a *Chlamydomonas* overview. *Genetics*, 179(1), pp.31-39.

Roach, T., Sedoud, A. and Krieger-Liszkay, A., 2013. Acetate in mixotrophic growth medium affects photosystem II in *Chlamydomonas reinhardtii* and protects against photoinhibition. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1827*(10), pp.1183-1190.

Roberts, A.G. and Kramer, D.M., 2001. Inhibitor "double occupancy" in the Q_o pocket of the chloroplast cytochrome b_6f complex. *Biochemistry*, *40*(45), pp.13407-13412.

Rochaix, J.D., 1995. *Chlamydomonas reinhardtii* as the photosynthetic yeast. *Annual review of genetics*, *29*(1), pp.209-230.

Rolland, N., Dorne, A.J., Amoroso, G., Sültemeyer, D.F., Joyard, J. and Rochaix, J.D., 1997. Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. *The EMBO Journal*, 16(22), pp.6713-6726.

Romero-Campero, F.J., Perez-Hurtado, I., Lucas-Reina, E., Romero, J.M. and Valverde, F., (2016). ChlamyNET: a *Chlamydomonas* gene co-expression network reveals global properties of the transcriptome and the early setup of key co-expression patterns in the green lineage. *BMC genomics*, *17*(1), p.227.

Ruban, A.V. and Johnson, M.P., 2009. Dynamics of higher plant photosystem cross-section associated with state transitions. *Photosynthesis research*, *99*(3), pp.173-183.

Sager, R. and Granick, S., 1953. Nutritional studies with *Chlamydomonas reinhardi*. *Annals of the New York Academy of Sciences*, *56*(5), pp.831-838.

Sager, R. and Palade, G.E., 1957. Structure and development of the chloroplast in *Chlamydomonas*: I. The normal green cell. *The Journal of Cell Biology*, *3*(3), pp.463-488.

Schirrmeister, B.E., Gugger, M. and Donoghue, P.C., 2015. Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology*, *58*(5), pp.769-785.

Schlarb-Ridley, B.G., Nimmo, R.H., Purton, S., Howe, C.J. and Bendall, D.S., 2006. Cytochrome c_{6A} is a funnel for thiol oxidation in the thylakoid lumen. *FEBS letters*, *580*(9), pp.2166-2169.

Schmidt, B., Ho, L. and Hogg, P.J., 2006. Allosteric disulfide bonds. *Biochemistry*, *45*(24), pp.7429-7433.

Schulz, H., Fabianek, R.A., Pellicioli, E.C., Hennecke, H. and Thöny-Meyer, L., 1999. Heme transfer to the heme chaperone CcmE during cytochrome c maturation requires the CcmC protein, which may function independently of the ABC-transporter CcmAB. *Proceedings of the National Academy of Sciences*, *96*(11), pp.6462-6467.

Schranz, M.E., Song, B.H., Windsor, A.J. and Mitchell-Olds, T., (2007). Comparative genomics in the Brassicaceae: a family-wide perspective. *Current opinion in plant biology*, *10*(2), pp.168-175.

Scranton, M.A., Ostrand, J.T., Fields, F.J. and Mayfield, S.P., 2015. *Chlamydomonas* as a model for biofuels and bio-products production. *The Plant Journal*, *82*(3), pp.523-531.

Seaver, L.C. and Imlay, J.A., 2001. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *Journal of bacteriology*, 183(24), pp.7173-7181.

Shao, N., Krieger-Liszkay, A., Schroda, M. and Beck, C.F., 2007. A reporter system for the individual detection of hydrogen peroxide and singlet oxygen: its use for the assay of reactive oxygen species produced *in vivo*. *The Plant Journal*, *50*(3), pp.475-487.

Shao, N., Duan, G.Y. and Bock, R., 2013. A mediator of singlet oxygen responses in *Chlamydomonas reinhardtii* and *Arabidopsis* identified by a luciferase-based genetic screen in algal cells. *The Plant Cell*, *25*(10), pp.4209-4226.

Shen, J.R. and Inoue, Y., 1993. Binding and functional properties of two new extrinsic components, cytochrome *c*-550 and a 12-kDa protein, in cyanobacterial photosystem II. *Biochemistry*, *32*(7), pp.1825-1832.

Slovacek, R.E., Mills, J.D. and Hind, G., 1978. The function of cyclic electron transport in photosynthesis. *FEBS Letters*, *87*(1), pp.73-76.

Spalding, M.H., 2007. Microalgal carbon-dioxide-concentrating mechanisms: *Chlamydomonas* inorganic carbon transporters. *Journal of experimental botany*, 59(7), pp.1463-1473.

Stauber, E.J. and Hippler, M., 2004. *Chlamydomonas reinhardtii* proteomics. *Plant Physiology and Biochemistry*, *42*(12), pp.989-1001.

Su, Y.Y.T. and Jirgensons, B., 1977. Further studies on detergent-induced conformational transitions in proteins: Circular dichroism of ovalbumin, bacterial α-amylase, papain, and β-lactoglobulin at various pH values. *Archives of biochemistry and biophysics*, *181*(1), pp.137-146.

Sueoka, N., 1960. Mitotic replication of deoxyribonucleic acid in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America*, *46*(1), p.83.

Suzuki, T., Yamasaki, K., Fujita, S., Oda, K., Iseki, M., Yoshida, K., Watanabe, M., Daiyasu, H., Toh, H., Asamizu, E. and Tabata, S., 2003. Archaeal-type rhodopsins in *Chlamydomonas*: model structure and intracellular localization. *Biochemical and biophysical research communications*, 301(3), pp.711-717.

Swidzinski, J.A., Leaver, C.J. and Sweetlove, L.J., 2004. A proteomic analysis of plant programmed cell death. *Phytochemistry*, *65*(12), pp.1829-1838.

Takahashi, T. and Watanabe, M., 1993. Photosynthesis modulates the sign of phototaxis of wild-type *Chlamydomonas reinhardtii*: Effects of red background illumination and 3-(3', 4'-dichlorophenyl)-1, 1-dimethylurea. *FEBS letters*, 336(3), pp.516-520.

Takahashi, M. and Asada, K., 1988. Superoxide production in aprotic interior of chloroplast thylakoids. *Archives of biochemistry and biophysics*, 267(2), pp.714-722.

Takahashi, S. and Murata, N., 2005. Interruption of the calvin cycle inhibits the repair of photosystem II from photodamage. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1708*(3), pp.352-361.

Takishita, K., Ishida, K.I. and Maruyama, T., 2003. An enigmatic GAPDH gene in the symbiotic dinoflagellate genus *Symbiodinium* and its related species (the order Suessiales): possible lateral gene transfer between two eukaryotic algae, dinoflagellate and euglenophyte. *Protist*, *154*(3-4), pp.443-454.

Teramoto, H., Nakamori, A., Minagawa, J. and Ono, T.A., 2002. Light-intensity-dependent expression of *Lhc* gene family encoding light-harvesting chlorophyll-a/b proteins of photosystem II in *Chlamydomonas reinhardtii*. *Plant Physiology*, *130*(1), pp.325-333.

Terauchi, A.M., Lu, S.F., Zaffagnini, M., Tappa, S., Hirasawa, M., Tripathy, J.N., Knaff, D.B., Farmer, P.J., Lemaire, S.D., Hase, T. and Merchant, S.S., 2009. Pattern of expression and substrate specificity of chloroplast ferredoxins from *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*, 284(38), pp.25867-25878.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y. and Stitt, M., 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*, 37(6), pp.914-939.
Thompson, J.D., Gibson, T.J. and Higgins, D.G., 2003. Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics*, (1), pp.2-3.

Tikkanen, M., Mekala, N.R. and Aro, E.M., 2014. Photosystem II photoinhibition-repair cycle protects Photosystem I from irreversible damage. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1837*(1), pp.210-215.

Tirumani, S., Kokkanti, M., Chaudhari, V., Shukla, M. and Rao, B.J., 2014. Regulation of CCM genes in *Chlamydomonas reinhardtii* during conditions of light–dark cycles in synchronous cultures. *Plant molecular biology*, *85*(3), pp.277-286.

Tolleter, D., Ghysels, B., Alric, J., Petroutsos, D., Tolstygina, I., Krawietz, D., Happe, T., Auroy, P., Adriano, J.M., Beyly, A. and Cuiné, S., 2011. Control of hydrogen photoproduction by the proton gradient generated by cyclic electron flow in *Chlamydomonas reinhardtii*. *The Plant Cell*, *23*(7), pp.2619-2630.

Ünlü, C., Drop, B., Croce, R. and van Amerongen, H., 2014. State transitions in *Chlamydomonas reinhardtii* strongly modulate the functional size of photosystem II but not of photosystem I. *Proceedings of the National Academy of Sciences*, *111*(9), pp.3460-3465.

Urzica, E.I., Adler, L.N., Page, M.D., Linster, C.L., Arbing, M.A., Casero, D., Pellegrini, M., Merchant, S.S. and Clarke, S.G., 2012. Impact of oxidative stress on ascorbate biosynthesis in *Chlamydomonas* via regulation of the *VTC2* gene encoding a GDP-L-galactose phosphorylase. *Journal of Biological Chemistry*, 287(17), pp.14234-14245.

Vallon, O., 1986. Lateral distribution of the main protein complexes of the photosynthetic apparatus in *Chlamydomonas reinhardtii* and in spinach: an immunocytochemical study using intact thylakoid membranes and a PSII enriched membrane preparation. *Photobiochem. Photobiophys.*, *12*, pp.203-220.

Vasudevan, R., 2019. Characterization of low molecular weight *c*-type cytochromes in cyanobacteria and plants. PhD thesis, University of Cambridge.

Virgin, I. and Andersson, B., 1988. Photosystem II disorganization and manganese release after photoinhibition of isolated spinach thylakoid membranes. *FEBS letters*, *233*(2), pp.408-412.

Wakabayashi, K.I. and King, S.M., 2006. Modulation of *Chlamydomonas reinhardtii* flagellar motility by redox poise. *J Cell Biol*, 173(5), pp.743-754.

Wakabayashi, K.I., Misawa, Y., Mochiji, S. and Kamiya, R., 2011. Reduction-oxidation poise regulates the sign of phototaxis in *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences*, 108(27), pp.11280-11284.

Walter, J.M., Coutinho, F.H., Dutilh, B.E., Swings, J., Thompson, F.L. and Thompson, C.C., 2017. Ecogenomics and taxonomy of Cyanobacteria phylum. *Frontiers in microbiology*, 8, p.2132.

Walters, R.G. and Horton, P., 1991. Resolution of components of non-photochemical chlorophyll fluorescence quenching in barley leaves. *Photosynthesis Research*, *27*(2), pp.121-133.

Wakao, S., Chin, B.L., Ledford, H.K., Dent, R.M., Casero, D., Pellegrini, M., Merchant, S.S. and Niyogi, K.K., 2014. Phosphoprotein SAK1 is a regulator of acclimation to singlet oxygen in *Chlamydomonas reinhardtii. Elife*, *3*, p.e02286.

Wang, Y., Stessman, D.J. and Spalding, M.H., 2015. The CO_2 concentrating mechanism and photosynthetic carbon assimilation in limiting CO_2: how *Chlamydomonas* works against the gradient. *The Plant Journal*, 82(3), pp.429-448.

Wang, L., Yamano, T., Takane, S., Niikawa, Y., Toyokawa, C., Ozawa, S.I., Tokutsu, R., Takahashi, Y., Minagawa, J., Kanesaki, Y. and Yoshikawa, H., 2016. Chloroplast-mediated regulation of CO_2-concentrating mechanism by Ca^{2+}-binding protein CAS in the green alga *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of S*ciences, 113(44), pp.12586-12591.

Wastl, J., Bendall, D.S. and Howe, C.J., 2002. Higher plants contain a modified cytochrome c_6. *Trends in plant science*, *7*(6), pp.244-245.

Wastl, J., Purton, S., Bendall, D.S. and Howe, C.J., 2004a. Two forms of cytochrome c_6 in a single eukaryote. *Trends in plant science*, *9*(10), pp.474-476.

Wastl, J., Molina-Heredia, F.P., Hervás, M., Navarro, J.A., De la Rosa, M.A., Bendall, D.S. and Howe, C.J., 2004b. Redox properties of *Arabidopsis* cytochrome c_6 are independent of the loop extension specific to higher plants. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1657*(2-3), pp.115-120.

Weigel, M., Pesaresi, P. and Leister, D., 2003a. Tracking the function of the cytochrome c_6-like protein in higher plants. *Trends in plant science*, *8*(11), pp.513-517.

Weigel, M., Varotto, C., Pesaresi, P., Finazzi, G., Rappaport, F., Salamini, F. and Leister, D., 2003b. Plastocyanin is indispensable for photosynthetic electron flow in *Arabidopsis thaliana*. *Journal of Biological Chemistry*, *278*(33), pp.31286-31289.

Wingler, A., Lea, P.J., Quick, W.P. and Leegood, R.C., 2000. Photorespiration: metabolic pathways and their role in stress protection. *Philosophical Transactions of the Royal Society of London*. *Series B: Biological Sciences*, 355(1402), pp.1517-1529.

Winkler, M., Hemschemeier, A., Jacobs, J., Stripp, S. and Happe, T., 2010. Multiple ferredoxin isoforms in *Chlamydomonas reinhardtii*–their role under stress conditions and biotechnological implications. *European journal of cell biology*, 89(12), pp.998-1004.

Wisecaver, J.H., Brosnahan, M.L. and Hackett, J.D., 2013. Horizontal gene transfer is a significant driver of gene innovation in dinoflagellates. *Genome biology and evolution*, *5*(12), pp.2368-2381.

Witman, G.B., 1993. *Chlamydomonas* phototaxis. *Trends in cell biology*, 3(11), pp.403-408.

Wood, P.M., 1978. Interchangeable copper and iron proteins in algal photosynthesis: studies on plastocyanin and cytochrome *c*-552 in *Chlamydomonas*. *European journal of biochemistry*, *87*(1), pp.9-19.

Worrall, J.A., Schlarb-Ridley, B.G., Reda, T., Marcaida, M.J., Moorlen, R.J., Wastl, J., Hirst, J., Bendall, D.S., Luisi, B.F. and Howe, C.J., 2007. Modulation of heme redox potential in the cytochrome c_6 family. *Journal of the American Chemical Society*, *129*(30), pp.9468-9475.

Worrall, J.A., Luisi, B.F., Schlarb-Ridley, B.G., Bendall, D.S. and Howe, C.J., 2008. Cytochrome c_{6A}: discovery, structure and properties responsible for its low haem redox potential. *Biochemical Society Transactions*, 36(6), pp.1175-1179.

Xiang, Y., Zhang, J. and Weeks, D.P., 2001. The *Cia5* gene controls formation of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, 98(9), pp.5341-5346.

Yamano, T., Miura, K. and Fukuzawa, H., 2008. Expression analysis of genes associated with the induction of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant physiology*, 147(1), pp.340-354.

Yan, N., Fan, C., Chen, Y. and Hu, Z., 2016. The potential for microalgae as bioreactors to produce pharmaceuticals. *International journal of molecular sciences*, *17*(6), p.962.

Yoshioka, S., Taniguchi, F., Miura, K., Inoue, T., Yamano, T. and Fukuzawa, H., 2004. The novel Myb transcription factor LCR1 regulates the CO_2-responsive gene *Cah1*, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. *The Plant Cell*, 16(6), pp.1466-1477.

Zatwarnicki, P., Barciszewski, J., Krzywda, S., Jaskolski, M., Kolesinski, P. and Szczepaniak, A., 2014. Cytochrome c_{6B} of *Synechococcus* sp. WH 8102–Crystal structure and basic properties of novel c_6-like family representative. *Biochemical and biophysical research communications*, *443*(4), pp.1131-1135.

Zito, F., Finazzi, G., Delosme, R., Nitschke, W., Picot, D. and Wollman, F.A., 1999. The Q_o site of cytochrome b_6f complexes controls the activation of the LHCII kinase. *The EMBO journal*, *18*(11), pp.2961-2969.

Žižková, E., Kubeš, M., Dobrev, P.I., Přibyl, P., Šimura, J., Zahajská, L., Záveská Drábková, L., Novák, O. and Motyka, V., 2016. Control of cytokinin and auxin homeostasis in cyanobacteria and algae. *Annals of botany*, 119(1), pp.151-166.

Appendix

Primers used in PCR

Target transcript	Forward primer	Reverse primer	Annealing temperature (°C)
CBLP	GTCTGTGGGACCTGAACAACC	TCCACAGCTTGATGGTCTTG	54
CYC4	AAAGAACGGCGTGGACAG	CACCTCCTCATCGCTCAGAC	54

Table S1: Sequences of the primers used in PCR. Annealing temperatures used in each PCR are also given.

Distances between cytochrome c_{6A} ORF (AT5G45040) and trypsin family protease (AT5G45030) or putative glutathione-S-transferase (AT5G45020)

Organism	Distance from trypsin family protease	Distance from putative glutathione-S-transferase
Arabidopsis thaliana	UTR overlaps	5 kbp
Arabidopsis halleri	UTR overlaps	Insufficient data
Arabidopsis lyrata	UTR overlaps	4 kbp
Boechera stricta	UTR overlaps	9 kbp
Brassica oleracea capitata	Not linked	1 kbp
Brassica rapa	Not linked	1 kbp
Capsella grandiflora	4 kbp	8 kbp
Capsella rubella	3 kbp	5 kbp
Eutrema salsugineum	UTR overlaps	3 kbp
Volvox carterae	Insufficient data	1.4 mbp
Chlamydomonas reinhardtii	Insufficient data	Not linked
Dunaliella salina	Insufficient data	Not linked
`Coccomyxa subellipsoidea	Insufficient data	Not linked
Micromonas pusilla CCMP1545	Insufficient data	Not linked
Micromonas SP RCC299	Insufficient data	Not linked
Ostreococcus lucimarinus	Insufficient data	71 kbp
Citrus sinensis	7 kbp	91 kbp
Citrus clementina	9 kbp	71 kbp
Carica papaya	29 kbp	89 kbp
Gossypium raimondii	85 kbp	Not linked
Theobroma cacao	15 kbp	24 kbp
Linum usitatissimum	7 mbp	12 kbp
Manihot esculenta	Insufficient data	Insufficient data
Populus trichocarpa	15 kbp	27 kbp
Ricinus communis	Insufficient data	Insufficient data
Salix purpurea	Insufficient data	Insufficient data
Eucalyptus grandis	Not linked	15 kbp

Vitis vinifera	9 kbp	22 kbp
Cucumis sativus	2 kbp	Not linked
Fragaria vesca	30 kbp	25 kbp
Glycine max	50 kbp	103 kbp
Malus domestica	Not linked	Not linked
Medicago truncatula	16 kbp	33 kbp
Phaseolus vulgaris	140 kbp	206 kbp
Prunus persica	8 kbp	17 kbp
Trifolium pratense	21 kbp	35 kbp
Aquilegia coerulea	85 kbp	54 kbp
Amaranthus hypochondriacus	Insufficient data	Insufficient data
Kalanchoe fedtschenkoi	Not linked	Not linked
Kalanchoe laxiflora	Not linked	Insufficient data
Daucus carota	1 kbp	4.8 mbp
Mimulus guttatus	2 kbp	8 kbp
Solanum lycopersicum	23 kbp	84 kbp
Solanum tuberosum	28 kbp	47 kbp
Panicum hallii	Not linked	Not linked
Panicum virgatum	Not linked	34.4 mbp
Setaria italica	Not linked	Not linked
Setaria viridis	Not linked	Not linked
Sorghum bicolor	Not linked	Not linked
Zea mays	Not linked	Not linked
Brachypodium distachhyon	Not linked	Not linked
Brachypodium stacei	Not linked	Not linked
Oryza sativa	Not linked	Not linked
Oropetium thomaeum	Not linked	Not linked
Marchantia polymorpha	Not linked	Not linked
Physcomitrella patens	Not linked	Not linked
Sphagnum fallax	Not linked	Not linked
Selaginella moellendorffii	Not linked	Not linked
Ananas comosus	Not linked	Not linked
Amborella trichopoda	222 kbp	Not linked
Musa acuminata	292 kbp	Not linked
Spirodela polyrhiza	Not linked	Not linked
Zostera marina	Not linked	Not linked

Table S2: Distances between cytochrome c_{6A} ORF (AT5G45040) and trypsin family protease (AT5G45030) or putative glutathione-S-transferase (AT5G45020) for each species listed in phyotozome, rounded to the nearest kbp (kilo-base pairs).

Quality control figures from RNA sequencing at Novogene

Sample Name	RNA Concentration (ng/µL)	Volume (µL)	Amount (µg)	OD260/280	OD260/230	RIN	Test Conclusion
Cw15 40 µE 1	182	17	3.094	2.22	1.94	8.9	Pass
Cw15 40 µE 2	184	16	2.944	2.09	2.3	8.8	Pass
Cw15 40 µE 3	206	17	3.502	2.15	2.06	9.1	Pass
Cw15 400 µE 1	246	15	3.69	2.28	2.02	9.1	Pass
Cw15 400 µE 2	208	16	3.328	2.26	1.43	9.1	Pass
Cw15 400 µE 3	190	15	2.85	2.32	2.26	8.9	Pass
Uex55 40 µE 1	190	17	3.23	2.02	2.02	8.9	Pass
Uex55 40 µE 2	214	16	3.424	2.18	1.78	8.7	Pass
Uex55 40 µE 3	156	15	2.34	2.11	2.11	9.2	Pass
Uex55 400 µE 1	238	17	4.046	2.16	1.86	8.1	Pass
Uex55 400 µE 2	244	18	4.392	2.1	2.18	8.5	Pass
Uex55 400 µE 3	214	16	3.424	2.06	2.38	8.5	Pass

Table S3: Table of quality control results for the extracted RNA as measured by Novogene using a Nanodrop. RIN stands for RNA integrity number, with an ideal value being close to 10.

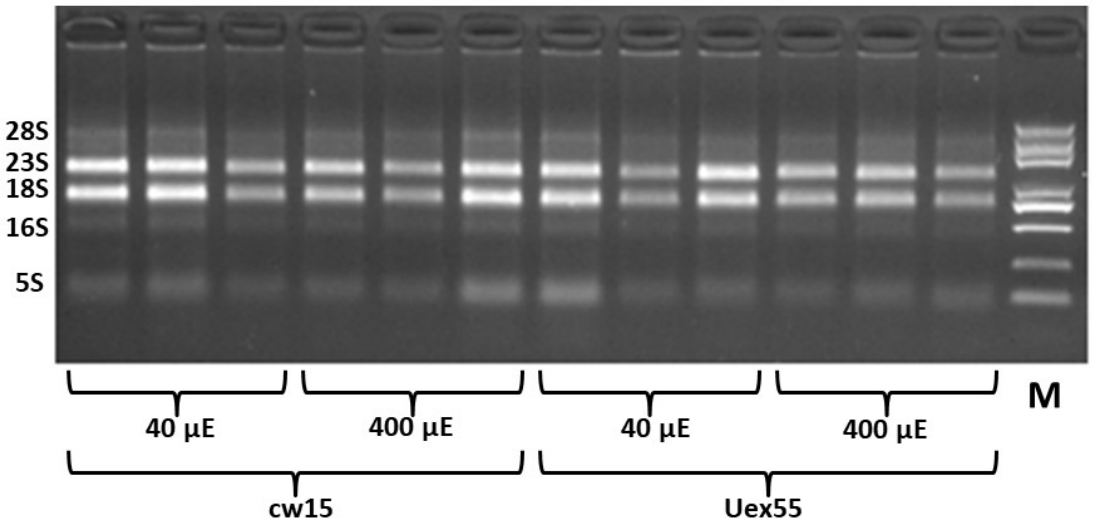
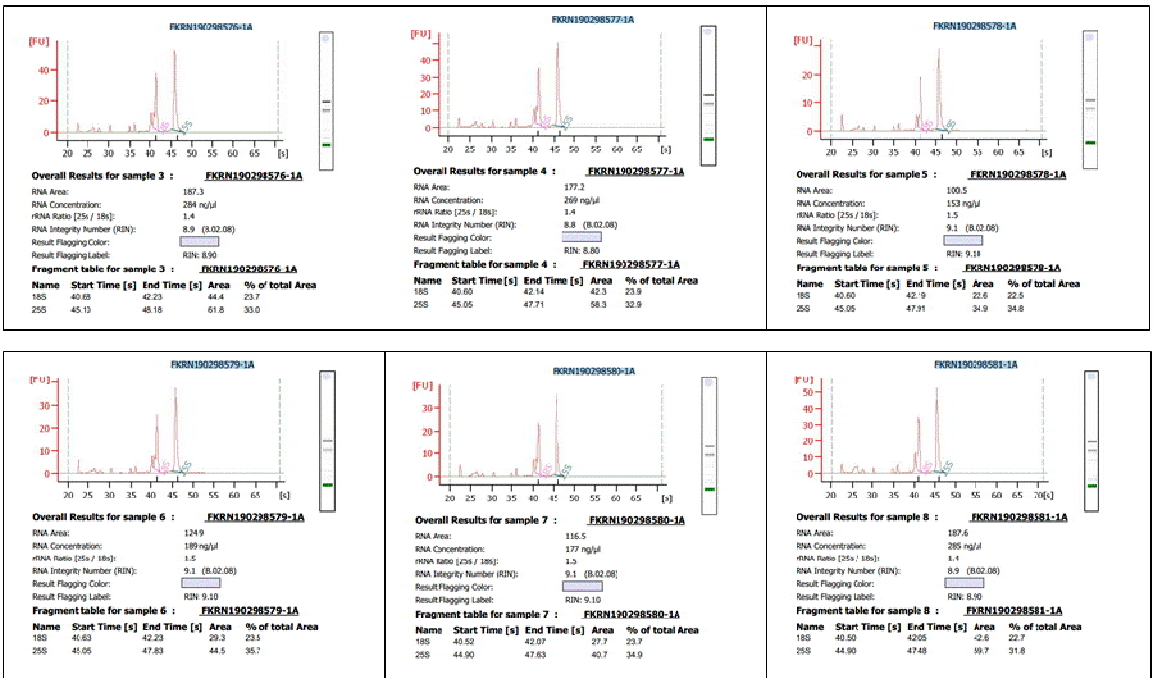


Figure S1: Electrophoresis gel of RNA samples run by Novogene. M represents the marker used (Trans2K Plus DNA Marker). The 28S, 23S, 18S, 16S and 5S bands have been labelled.



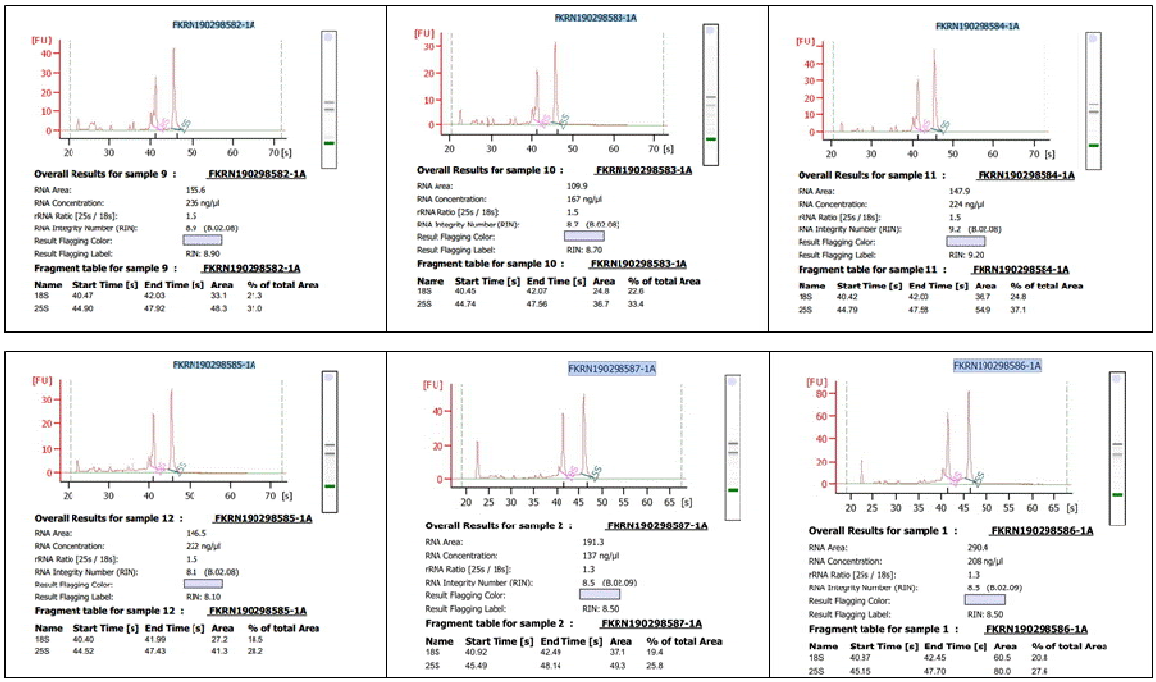


Figure S2: Sample analysis using the Agilent 2100 bioanalyser. The graph shows the fluorescence units (FU) for each of the different sized RNAs in the sample as separated by capillary chromatography (bar on far right). The 18S and 25S ribosomal RNA peaks have been labelled in pink and green respectively. The areas of these peaks were used to calculate various properties of the sample, which are listed. The three replicates samples are shown on each row, and the rows are in the order: cw15 LL, cw15 HL, Uex55 LL and Uex55 HL. Data analysis provided by Novogene.

Sample name	Raw reads	Clean Reads	Raw data	Clean data	Effective (%)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
Cw15 40 µE 1	31,014,202	30,483,824	9,304,260,600	9,145,147,200	98.29	0.02	98.32	95.33	63.54
Cw15 40 µE 2	27,423,400	27,138,224	8,227,020,000	8,141,467,200	98.96	0.02	98.31	95.32	63.58
Cw15 40 µE 3	33,216,267	32,784,840	9,964,880,100	9,835,452,000	98.7	0.02	98.32	95.34	63.51
Cw15 400 µE	31,735,452	31,380,229	9,520,635,600	9,414,068,700	98.88	0.02	98.25	95.19	63.48

1									
Cw15 400 µE 2	27,704,887	27,345,050	8,311,466,100	8,203,515,000	98.7	0.02	98.47	95.68	63.73
Cw15 400 µE 3	32,486,143	32,115,626	9,745,842,900	9,634,687,800	98.86	0.02	98.23	95.12	63.66
Uex55 40 µE 1	28,521,652	28,348,445	8,556,495,600	8,504,533,500	99.39	0.02	98.22	95.1	63.7
Uex55 40 µE 2	34,125,441	33,852,747	10,237,632,300	10,155,824,100	99.2	0.02	98.28	95.22	63.66
Uex55 40 µE 3	34,351,418	34,043,477	10,305,425,400	10,213,043,100	99.1	0.02	98.34	95.39	63.69
Uex55 400 µE 1	30,472,549	30,198,621	9,141,764,700	9,059,586,300	99.1	0.02	98.36	95.41	63.55
Uex55 400 µE 2	29759206	29504324	8927761800	8851297200	99.14	0.02	98.32	95.29	63.47
Uex55 400 µE 3	29008116	28712270	8702434800	8613681000	98.98	0.02	98.38	95.46	63.69

Table S4: Summary of the Novogene quality control of the RNA sequencing reaction for each sample. Q20 and Q30 represent sequences that have an error rate of 0.01 and 0.001 % respectively. GC refers to the percentage of the bases guanine and cytosine in the sequences.

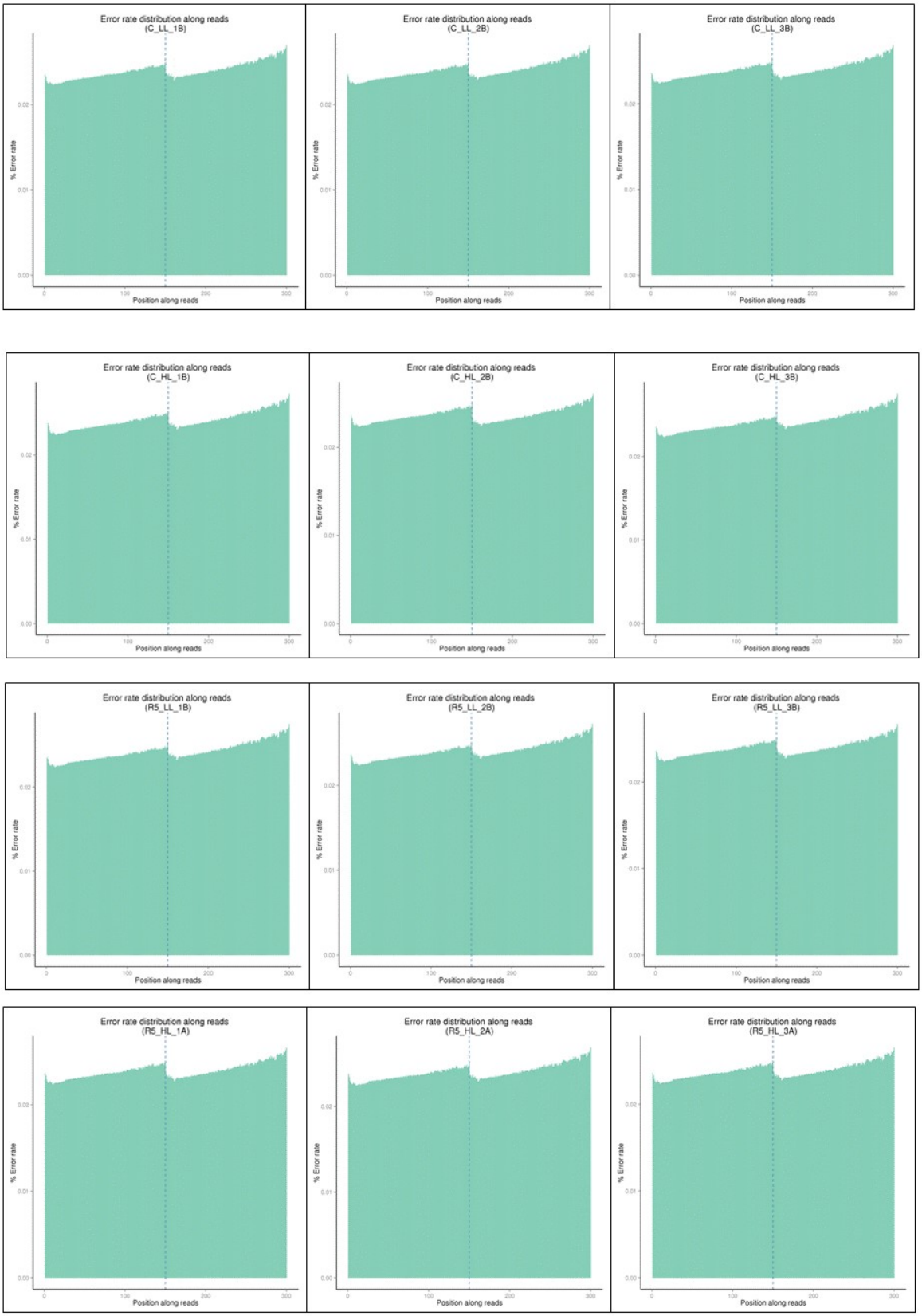


Figure S3: The distribution of the error rate (e) along the RNA sequencing reads. The three replicates samples are shown on each row, and the rows are in the order: cw15 LL, cw15 HL, Uex55 LL and Uex55 HL. Data analysis provided by Novogene.

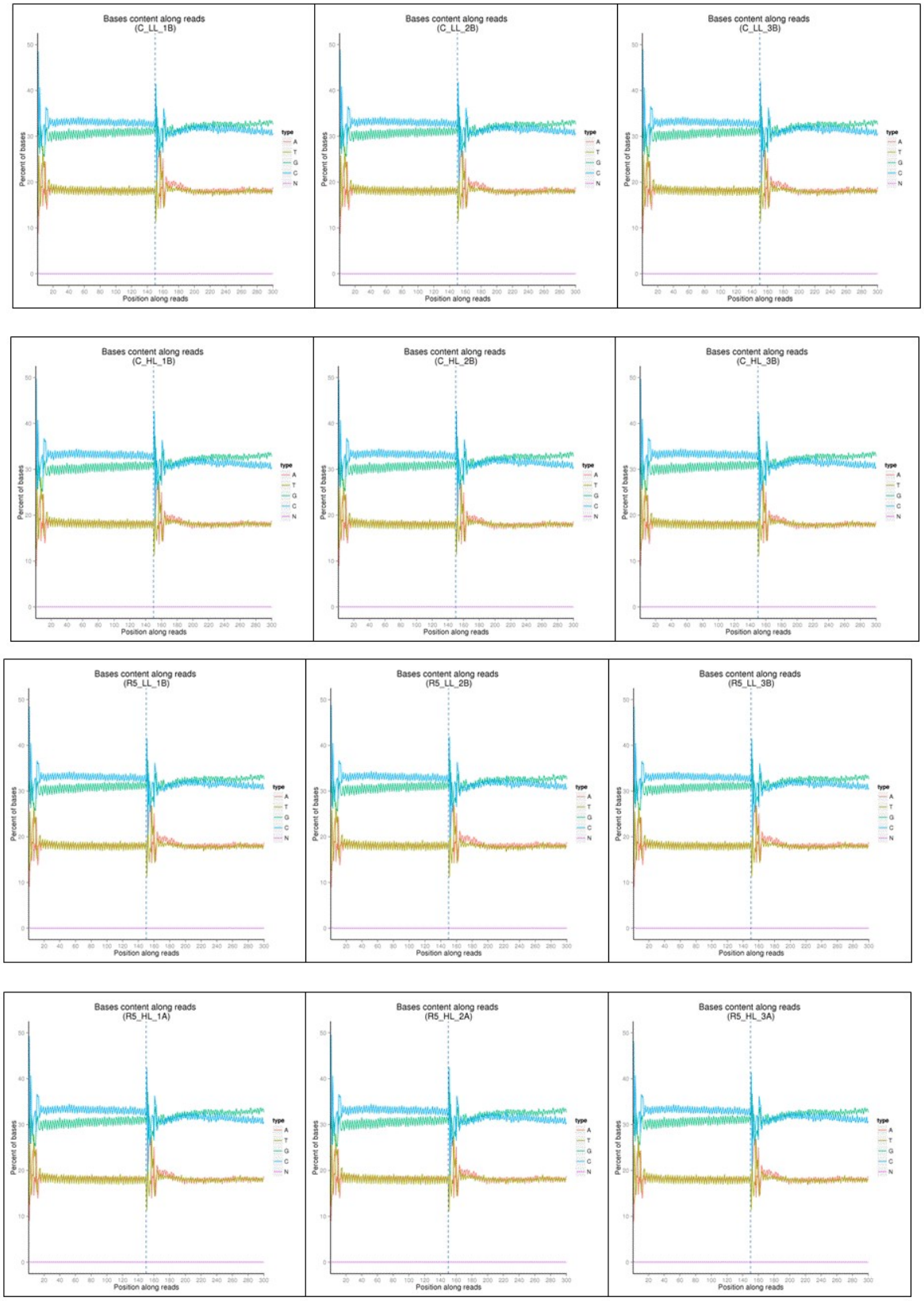
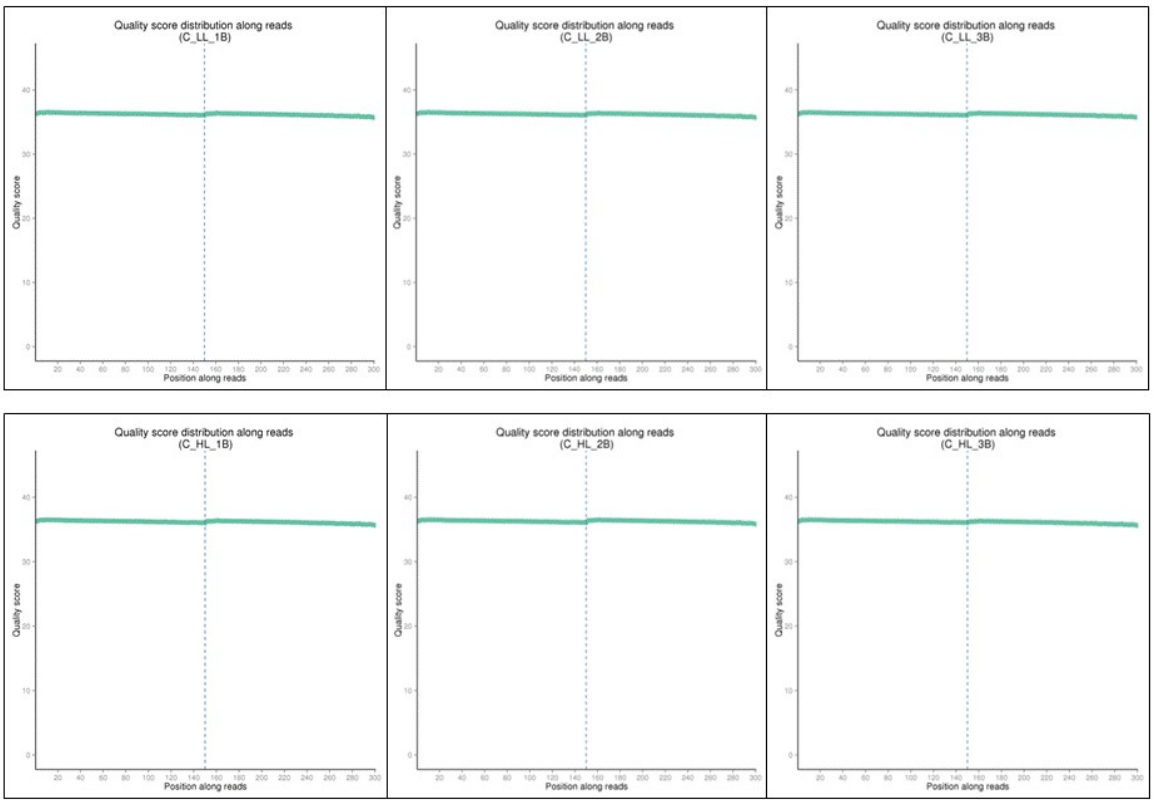


Figure S4: The distribution of base content along the RNA sequencing reads shows the relative frequency of A (red), T (yellow), G (green), C (blue) and N (pink; representing unidentifiable bases).

The three replicates samples are shown on each row, and the rows are in the order: cw15 LL, cw15 HL, Uex55 LL and Uex55 HL. Data analysis provided by Novogene.

Figure S5: The percentage of raw reads classified as low quality (orange), containing too many unidentified bases (yellow), containing adapters (green) or 'clean' and suitable for further analysis (purple). The three replicates samples are shown on each row, and the rows are in the order: cw15 LL, cw15 HL, Uex55 LL and Uex55 HL. Data analysis provided by Novogene.



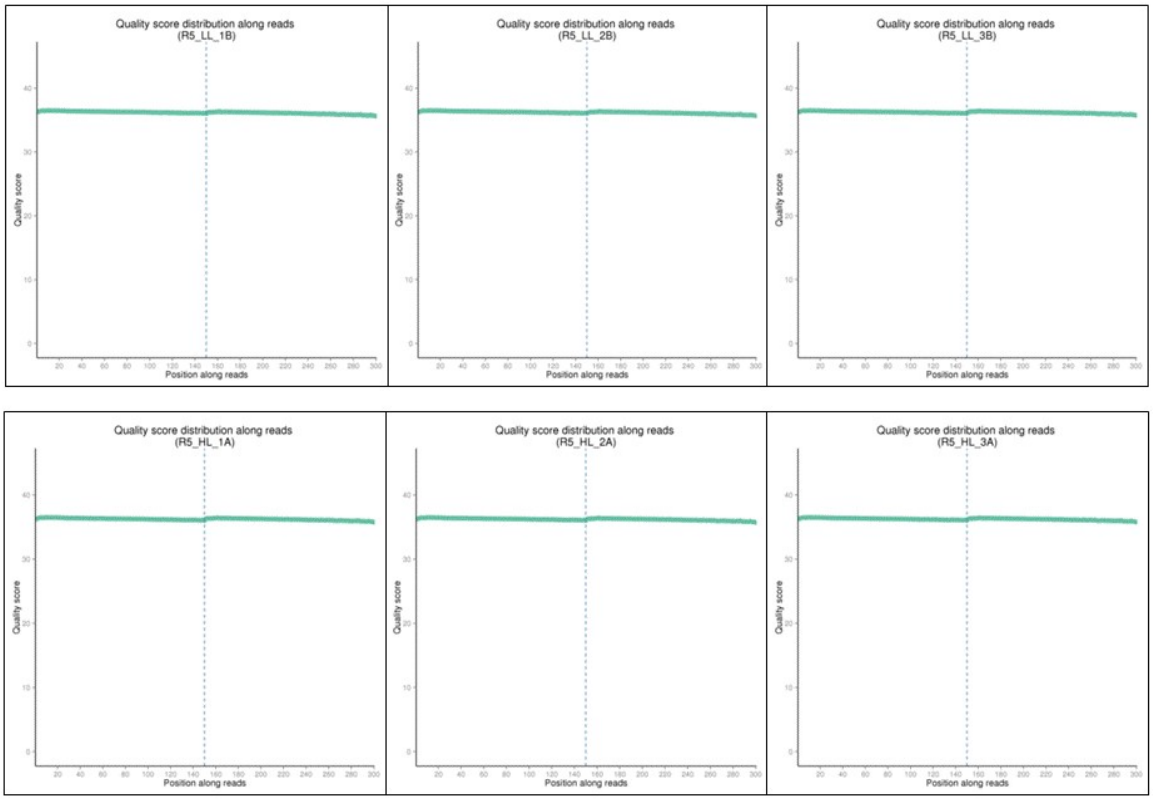


Figure S6: The distribution of sequence quality shows how the quality score (-10xlog_{10}(e)) changes along the RNA sequencing reads. The three replicates samples are shown on each row, and the rows are in the order: cw15 LL, cw15 HL, Uex55 LL and Uex55 HL. Data analysis provided by Novogene.